

**MORGAN & MORGAN COMPLEX
LITIGATION GROUP**

Michael F. Ram (SBN 104805)
mram@forthepeople.com
711 Van Ness Ave, Suite 500
San Francisco, CA 94102
Tel.: (415) 358-6913

**MORGAN & MORGAN COMPLEX
LITIGATION GROUP**

John A. Yanchunis (*pro hac vice forthcoming*)
jyanchunis@forthepeople.com
Ryan J. McGee (*pro hac vice forthcoming*)
rmcgee@forthepeople.com
201 N. Franklin St., 7th Floor
Tampa, FL 33602
Tel.: (813) 223-5505

Counsel for Plaintiffs and the Proposed Classes

**UNITED STATES DISTRICT COURT
NORTHERN DISTRICT OF CALIFORNIA**

PLAINTIFFS A.T., J.H., individually, and on
behalf of all others similarly situated,

Plaintiffs,

vs.

OPENAI LP, OPENAI INCORPORATED,
OPENAI GP, LLC, OPENAI STARTUP FUND
I, LP, OPENAI STARTUP FUND GP I, LLC,
OPENAI STARTUP FUND MANAGEMENT
LLC, MICROSOFT CORPORATION and DOES
1 through 20, inclusive,

Defendants.

Case No.:

CLASS ACTION COMPLAINT

1. VIOLATION OF ELECTRONIC COMMUNICATIONS PRIVACY ACT, 18 U.S.C. §§ 2510, *et seq.*
2. VIOLATION OF THE COMPUTER FRAUD AND ABUSE ACT, 18 U.S.C. § 1030
3. VIOLATION OF THE CALIFORNIA INVASION OF PRIVACY ACT (“CIPA”), CAL. PENAL CODE § 631
4. VIOLATION OF CALIFORNIA UNFAIR COMPETITION LAW, BUSINESS AND PROFESSIONS CODE §§ 17200, *et seq.*
5. NEGLIGENCE
6. INVASION OF PRIVACY
7. INTRUSION UPON SECLUSION
8. LARCENY/RECEIPT OF STOLEN PROPERTY
9. CONVERSION
10. UNJUST ENRICHMENT
11. NEW YORK GENERAL BUSINESS LAW §§ 349, *et seq.*

DEMAND FOR JURY TRIAL

TABLE OF CONTENTS

1

2 INTRODUCTION1

3 PARTIES8

4 JURISDICTION AND VENUE13

5 FACTUAL BACKGROUND14

6 I. DEVELOPMENT OF ARTIFICIAL INTELLIGENCE IN THE U.S.....14

7 A. OpenAI: From Open Nonprofit to Profit-Driven \$29B Commercial Partner of Tech Giant

8 Microsoft.....14

9 B. OpenAI’s Products.....18

10 C. ChatGPT’s Development Depends on Secret Web-Scraping20

11 D. ChatGPT Training on Users of Defendants’ Programs and Applications.25

12 E. Microsoft Pushes OpenAI’s Economic Dependence Model27

13 II. Risks from Unchecked AI Proliferation30

14 A. The International Community Agrees that Unchecked & Lawless AI Proliferation Poses

15 an Existential Threat30

16 B. Overview of Risks.....35

17 1. *Massive Privacy Violations*.....35

18 2. *AI-Fueled Misinformation Campaigns, Targeted Attacks, Sex Crimes, and Bias*.....39

19 3. *Hypercharged Malware Creation*.....43

20 4. *Autonomous Weapons*.....46

21 C. Opportunity on the Other Side.....47

22 III. DEFENDANTS’ CONDUCT VIOLATES ESTABLISHED PROPERTY AND PRIVACY

23 RIGHTS50

24 A. Defendants’ Web-Scraping Theft50

25 B. Defendants’ Web Scraping Violated Plaintiff’s Property Interests52

26 C. Defendants’ Web Scraping Violated Plaintiffs’ Privacy Interests.....54

27 D. Defendants’ Business Practices are Offensive to Reasonable People and Ignore

28 Increasingly Clear Warnings from Regulators57

1 E. Defendants’ Theft of User Data in Excess of Reasonable Consent.....60

2 1. *OpenAI’s disclosures are not conspicuous*.....63

3 2. *Defendants’ Use of Consumer Data Far Exceeds Industry Standards and their Own*

4 *Representations*.....64

5 CLASS ALLEGATIONS67

6 CALIFORNIA LAW SHOULD APPLY TO OUT-OF-STATE PLAINTIFF’S & CLASS

7 MEMBERS’ NON-STATUTORY CLAIMS74

8 COUNT ONE: VIOLATION OF ELECTRONIC COMMUNICATIONS PRIVACY ACT, 18

9 U.S.C. § 2510, *et seq.*.....75

10 I. Interception of Communications Between ChatGPT API Class Members which occurred on

11 Third-Party Websites, Platforms, Applications, Programs which have integrated ChatGPT API.

12 [Microsoft User Class is Excluded]79

13 II. Microsoft’s Interception of Communications Between ChatGPT Class Members81

14 III. Defendant Open AI’s Interception of Microsoft User Class Members which occurred on

15 Microsoft’s Websites, Platforms, Applications, Programs which have integrated ChatGPT. .83

16 COUNT TWO: VIOLATION OF THE COMPUTER FRAUD AND ABUSE ACT, 18 U.S.C. §

17 1030.....85

18 COUNT THREE: VIOLATION OF THE CALIFORNIA INVASION OF PRIVACY ACT

19 (“CIPA”), CAL. PENAL CODE § 631, *et seq.*87

20 I. Defendants’ Interception of Communications of ChatGPT API Class Members which occurred

21 on Third-Party Websites, Platforms, Applications, Programs which have integrated ChatGPT

22 API. [Microsoft User Subclass is Excluded]90

23 II. Microsoft’s Interception of ChatGPT User Class Members’ Communications on ChatGPT ..92

24 III. Defendant Open AI’s Interception of Microsoft User Class Members which occurred on

25 Microsoft’s Websites, Platforms, Applications, Programs which have integrated ChatGPT. .94

26 COUNT FOUR: VIOLATION OF CALIFORNIA UNFAIR COMPETITION LAW (Cal. Bus. &

27 Prof. Code §§ 17200, *et seq.*).....96

28 I. Unlawful96

1 II. Unfair101
2 COUNT FIVE: NEGLIGENCE.....104
3 COUNT SIX: INVASION OF PRIVACY105
4 COUNT SEVEN: INTRUSION UPON SECLUSION108
5 COUNT EIGHT: LARCENY/RECEIPT OF STOLEN PROPERTY CAL. PENAL CODE §§
6 496(a) and 496(c)109
7 A. Defendants’ Taking of Individual’s Private Information to Train Their AI Violated
8 Plaintiffs’ Property Interests109
9 B. Tracking, Collecting, and Sharing Private Information Without Consent.....110
10 COUNT NINE: CONVERSION111
11 COUNT TEN: UNJUST ENRICHMENT112
12 COUNT ELEVEN: NEW YORK GENERAL BUSINESS LAW §§ 349, *et seq.*113
13 PRAYER FOR RELIEF.....114
14 JURY TRIAL DEMANDED117

15
16
17
18
19
20
21
22
23
24
25
26
27
28

1 Plaintiffs A.T. and J.H. (hereinafter “Plaintiffs”),¹ individually and on behalf of all others
 2 similarly situated, bring this action against Defendants OpenAI, OpenAI Incorporated, OpenAI GP
 3 LLC, OpenAI Startup Fund I, LP, OpenAI Startup Fund GP I, LLC, and Microsoft Corporation
 4 (collectively, “Defendants”). Plaintiff’s allegations are based upon personal knowledge as to
 5 themselves and their own acts, and upon information and belief as to all other matters based on the
 6 investigation conducted by and through Plaintiff’s attorneys.

7 INTRODUCTION

8 1. On October 19, 2016, University of Cambridge Professor of Theoretical Physics
 9 Stephen Hawking predicted, “Success in creating AI could be the biggest event in the history of our
 10 civilization. But it could also be the last, unless we learn how to avoid the risks.”² Professor Hawking
 11 described a future in which humanity would choose to either harness the huge potential benefits or
 12 succumb to the dangers of AI, emphasizing “the rise of powerful AI will be either the best or the
 13 worst thing ever to happen to humanity.”

14 2. The future Professor Hawking predicted has arrived in just seven short years. Using
 15 stolen and misappropriated personal information at scale, Defendants have created powerful and
 16 wildly profitable AI and released it into the world without regard for the risks. In so doing,
 17 Defendants have created an AI arms race in which Defendants and other Big Tech companies are
 18 onboarding society into a plane that over half of the surveyed AI experts believe has at least a 10%
 19 chance of crashing and killing everyone on board.³ Humanity is now faced with the two Frostian
 20

21 ¹ Plaintiffs respectfully request that the Court permit them to keep their identity private as
 22 Plaintiffs aim to avoid intrusive scrutiny as well as any potentially dangerous backlash. Indeed,
 23 plaintiffs in other lawsuits against the same defendant entities have received many troubling and
 24 violent threats, including death threats, marking a severe infringement of personal safety.
 Accordingly, opting for privacy is a critical measure to avoid unwarranted negative attention as
 well as potential harm. Plaintiffs will file a motion to proceed pseudonymously, if required. *See*
 Victoria Hudgins, *GitHub and Openai Plaintiffs Seek Anonymity amid Slurs and Death Threats*,
 GLOB. DATA REV. (Mar. 15, 2023), globaldatareview.com/article/github-and-openai-plaintiffs-see-anonymity-amid-slurs-and-death-threats.

25 ² Cambridge University, *The Best or Worst Thing to Happen to Humanity*, YOUTUBE (Oct. 19,
 26 2016), https://www.youtube.com/watch?v=_5XvDCjrdXs&t=1s.

27 ³ Yuval Harari et al., *You Can Have the Blue Pill or the Red Pill, and We’re Out of Blue Pills*,
 THE N.Y. TIMES (Mar. 24, 2023), [https://www.nytimes.com/2023/03/24/opinion/yuval-harari-ai-
 chatgpt.html](https://www.nytimes.com/2023/03/24/opinion/yuval-harari-ai-chatgpt.html) (“[O]ver 700 top academics and researchers behind the leading artificial intelligence
 28 companies were asked in a survey about future A.I. risk. Half of those surveyed stated that there
 was a 10 percent or greater chance of human extinction (or similarly permanent and severe
 disempowerment) from future A.I. systems.”).

1 roads Professor Hawking predicted we would have to choose between: One leads to sustainability,
2 security, and prosperity; the other leads to civilizational collapse.

3 3. This class action lawsuit arises from Defendants’ unlawful and harmful conduct in
4 developing, marketing, and operating their AI products, including ChatGPT-3.5, ChatGPT-4.0,⁴
5 Dall-E, and Vall-E (the “**Products**”), which use stolen private information, including personally
6 identifiable information, from hundreds of millions of internet users, including children of all ages,
7 without their informed consent or knowledge. Furthermore, Defendants continue to unlawfully
8 collect and feed additional personal data from millions of unsuspecting consumers worldwide, far
9 in excess of any reasonably authorized use, in order to continue developing and training the
10 Products.

11 4. Defendants’ disregard for privacy laws is matched only by their disregard for the
12 potentially catastrophic risk to humanity. Emblematic of both the ultimate risk—and Defendants’
13 open disregard—is this statement from Defendant OpenAI’s CEO Sam Altman: “AI will probably
14 most likely lead to the end of the world, but in the meantime, there’ll be great companies.”⁵

15 5. Defendants’ Products, and the technology on which they are built, undoubtedly have
16 the potential to do much good in the world, like aiding life-saving scientific research and ushering
17 in discoveries that can improve the lives of everyday Americans. With that potential in mind,
18 Defendant OpenAI was originally founded as a nonprofit research organization with a single
19 mission: to create and ensure artificial intelligence would be used for the benefit of humanity. But
20 in 2019, OpenAI abruptly restructured itself, developing a for-profit business that would pursue
21 commercial opportunities of staggering scale.

22 6. As a result of the restructuring, OpenAI abandoned its original goals and principles,
23 electing instead to pursue profit at the expense of privacy, security, and ethics. It doubled down on

24 ⁴ ChatGPT is referred to herein as inclusive of both ChatGPT-3.5, ChatGPT-4, and any other
25 versions of ChatGPT. The term “ChatGPT Plug-In” encompasses GPT-3.5, GPT-4, and any
26 additional extensions that have been incorporated into Microsoft’s and third-party platforms,
websites, applications, programs, or systems.

27 ⁵ Matt Weinberger, *Head of Silicon Valley’s Most Important Startup Firm Says We’re in A ‘Mega*
Bubble’ That Won’t Last, BUS. INSIDER (June 4, 2015), [https://www.businessinsider.com/sam-](https://www.businessinsider.com/sam-altman-y-combinator-talks-mega-bubble-nuclear-power-and-more-2015-6?r=US)
28 [altman-y-combinator-talks-mega-bubble-nuclear-power-and-more-2015-6?r=US](https://www.businessinsider.com/sam-altman-y-combinator-talks-mega-bubble-nuclear-power-and-more-2015-6?r=US); David Wallace-
Wells, *A.I. Is Being Built by People Who Think It Might Destroy Us*, THE N.Y. TIMES (Mar. 27,
2023), <https://www.nytimes.com/2023/03/27/opinion/ai-chatgpt-chatbots.html>.

1 a strategy to secretly harvest massive amounts of personal data from the internet, including private
2 information and private conversations, medical data, information about children—essentially every
3 piece of data exchanged on the internet it could take—without notice to the owners or users of such
4 data, much less with anyone’s permission.

5 7. Without this unprecedented theft of private and copyrighted information belonging to
6 real people, communicated to unique communities, for specific purposes, targeting specific
7 audiences, the Products would not be the multi-billion-dollar business they are today. OpenAI used
8 the stolen data to train and develop the Products utilizing large language models (LLMs) and deep
9 language algorithms to analyze and generate human-like language that can be used for a wide range
10 of applications, including chatbots, language translation, text generation, and more. Defendants’
11 Products’ sophisticated natural language processing capabilities allow them to, among other things,
12 carry on human-like conversations with users, answer questions, provide information, generate next
13 text on demand, create art, and connect emotionally with people, all like a “real” human.

14 8. Once trained on stolen data, Defendants saw the immediate profit potential and rushed
15 the Products to market without implementing proper safeguards or controls to ensure that they
16 would not produce or support harmful or malicious content and conduct that could further violate
17 the law, infringe rights, and endanger lives. Without these safeguards, the Products have already
18 demonstrated their ability to harm humans, in real ways.

19 9. A nontrivial number of experts claim the risks to humanity presented by the Products
20 outweigh even those of the Manhattan Project’s development of nuclear weapons. Historically, the
21 unchecked release of new technologies without proper safeguards and regulations has caused
22 chaos.⁶ Now again, we face imminent and unreasonable risks of the very fabric of our society

23 ⁶ Bill Kovarik, *A Century of Tragedy: How the Car and Gas Industry Knew About The Health*
24 *Risks of Leaded Fuel But Sold it For 100 Years Anyway*, THE CONVERSATION (Dec. 8, 2021),
25 [https://theconversation.com/a-century-of-tragedy-how-the-car-and-gas-industry-knew-about-the-](https://theconversation.com/a-century-of-tragedy-how-the-car-and-gas-industry-knew-about-the-health-risks-of-leaded-fuel-but-sold-it-for-100-years-anyway-173395)
26 [health-risks-of-leaded-fuel-but-sold-it-for-100-years-anyway-173395](https://theconversation.com/a-century-of-tragedy-how-the-car-and-gas-industry-knew-about-the-health-risks-of-leaded-fuel-but-sold-it-for-100-years-anyway-173395) (1920s invention of leaded
27 gasoline, initially thought of as a technological breakthrough, resulted in serious health and
28 environmental consequences, such as lead poisoning and soil contamination); James H. Kim &
Anthony R. Scialli, *Thalidomide: The Tragedy of Birth Defects and the Effective Treatment of*
Disease, 122 TOXICOLOGICAL SCI. 1, 1 (2011) (Development of thalidomide in the 1950s and 60s,

1 unraveling, at the hands of profit-driven, multibillion-dollar corporations.

2 10. Powerful companies, armed with unparalleled and highly concentrated technological
3 capabilities, have recklessly raced to release AI technology with disregard for the catastrophic risk
4 to humanity in the name of “technological advancement.” As the National Security Commission
5 noted in its Final Report on AI, “the U.S. government is a long way from being ‘AI-ready.’”⁷

6 11. Experts believe that without immediate legal intervention this will lead to scenarios
7 where AI can act against human interests and values, exploit human beings⁸ without regard for their
8 well-being or consent, and/or even decide to eliminate the human species as a threat to its goals. As
9 Geoffrey Everest Hinton—the seminal figure in the development of the technology on which the
10 Products run—put it: “The alarm bell I’m ringing has to do with the existential threat of them taking
11 control... I used to think it was a long way off, but now I think it’s serious and fairly close.”⁹ He is

12
13 thought to be the miraculous solution to nausea, led to widespread birth defects in babies whose
14 mothers had taken the drug); PWJ Bartrip, *History of Asbestos Related Disease*, 80
15 POSTGRADUATE MED. J. 72, 72-5 (Feb. 2004) (Introduction of asbestos in the early 20th century,
16 later found to cause lung cancer and other serious health problems, leading to bans and strict
17 regulation); Jason Von Meding, *Agent Orange, Exposed: How U.S. Chemical Warfare in Vietnam*
18 *Unleashed a Slow-Moving Disaster*, THE CONVERSATION (Oct. 3, 2017),
19 [https://theconversation.com/agent-orange-exposed-how-u-s-chemical-warfare-in-vietnam-](https://theconversation.com/agent-orange-exposed-how-u-s-chemical-warfare-in-vietnam-unleashed-a-slow-moving-disaster-84572)
20 [unleashed-a-slow-moving-disaster-84572](https://theconversation.com/agent-orange-exposed-how-u-s-chemical-warfare-in-vietnam-unleashed-a-slow-moving-disaster-84572) (The U.S. military’s deployment of over 45 million
21 liters of toxic chemical Agent Orange unleashed a health and ecological disaster, causing life-
22 threatening birth defects in children and destroying forests and habitats across Vietnam).

23 ⁷ 2021 Final Report, NAT. SEC. COMM. ON A.I., www.nsc.gov/2021-final-report/ (last visited
24 June 27, 2023).

25 ⁸ CAPTCHAs allow websites to determine whether users are human or bots. Traditionally,
26 CAPTCHAs involve “puzzles or image recognition tasks that are challenging for automated
27 programs but straightforward for humans to solve.” These tests are used widely across the web to
28 prevent bots from spamming websites, creating fake accounts, or scraping content. In one recent,
troubling incident, ChatGPT 4 evaded CAPTCHA safeguards by hiring a human worker from
TaskRabbit, a crowdsourcing platform, to solve CAPTCHAs on its behalf, tricking the worker
into believing it was a human with visual impairment. See *ChatGPT 4 Hires a TaskRabbit and*
Tricks Them into Completing a CAPTCHA, INTERESTING SOUP (Mar. 15, 2023),
<https://interestingsoup.com/gpt4-requests-a-taskrabbit-to-solve-captcha-for-it/>; Beatrice Nolan,
The Latest Version of ChatGPT Told a Taskrabbit Worker it was Visually Impaired to Get Help
Solving a CAPTCHA, OpenAI Test Shows, BUS. INSIDER (Mar. 16, 2023),
[https://www.businessinsider.com/gpt4-openai-chatgpt-taskrabbit-tricked-solve-captcha-test-2023-](https://www.businessinsider.com/gpt4-openai-chatgpt-taskrabbit-tricked-solve-captcha-test-2023-3)
3.

⁹ Craig S. Smith, *Geoff Hinton, AI’s Most Famous Researcher, Warns of ‘Existential Threat’*
From AI, FORBES (May 4, 2023), [https://www.forbes.com/sites/craigsmith/2023/05/04/geoff-](https://www.forbes.com/sites/craigsmith/2023/05/04/geoff-hinton-ais-most-famous-researcher-warns-of-existential-threat/?sh=1ffcd7a65215)
[hinton-ais-most-famous-researcher-warns-of-existential-threat/?sh=1ffcd7a65215](https://www.forbes.com/sites/craigsmith/2023/05/04/geoff-hinton-ais-most-famous-researcher-warns-of-existential-threat/?sh=1ffcd7a65215).

1 not alone.¹⁰

2 12. While the downsides are nearly unimaginable, the upsides are similarly archetype-
3 shattering. Defendant OpenAI’s technology is already valued at tens of billions of dollars, and its
4 reach into every public and private industry continues apace. The Products only reached the level
5 of sophistication they have today due to training on stolen, misappropriated data, and Defendants
6 continue to misappropriate data, scraping from the internet without any notice or consent, as well
7 as taking personal information from the Products’ 100+ million registered users without their full
8 knowledge and consent.

9 13. Additionally, the Products are increasingly being incorporated into an ever-expanding
10 roster of applications and websites, through either API or plug-ins.¹¹ Through integration of
11 Defendants’ AI in nearly every possible product and industry, Defendants created and continue to
12 create economic dependency within our society, deploying the tech directly into the hands of society
13 and embedding it into the fundamental infrastructure as quickly as possible. As posed by Center for
14 Humane Technology Cofounders Tristan Harris and Aza Raskin in their carefully crafted critique
15 of the rapid deployment of AI, “Do you think that once [these industries] discover some problem
16 that they [will] just withdraw or retract it from society? No, increasingly, the government, militaries
17 [and others], are rapidly building their whole next systems and raising venture capital to build on
18 top of this layer of society... *That’s not testing it with society, that is onboarding humanity onto*
19 *an untested plane... It’s one thing to test, it’s another thing to create economic dependency.*”¹²

20 14. The head of the alignment team and safety at Open AI directly acknowledges these
21 risks, postulating, “before we scramble to deeply integrate large language models everywhere in the
22 economy, can we pause and think whether it is wise to do so? This is quite immature technology,

23 ¹⁰ James Vincent, *Top AI Researchers and CEOs Warn Against ‘Risk of Extinction’ in 22 Word*
24 *Statement*, THE VERGE (May 30, 2023), [https://www.theverge.com/2023/5/30/23742005/ai-risk-](https://www.theverge.com/2023/5/30/23742005/ai-risk-warning-22-word-statement-google-deepmind-openai)
25 [warning-22-word-statement-google-deepmind-openai](https://www.theverge.com/2023/5/30/23742005/ai-risk-warning-22-word-statement-google-deepmind-openai).

26 ¹¹ *Here are the Companies Using ChatGPT*, GADGETS NOW (Mar. 17, 2023),
27 [https://www.gadgetsnow.com/slideshows/here-are-the-companies-using-](https://www.gadgetsnow.com/slideshows/here-are-the-companies-using-chatgpt/photolist/98735402.cms)
28 [chatgpt/photolist/98735402.cms](https://www.gadgetsnow.com/slideshows/here-are-the-companies-using-chatgpt/photolist/98735402.cms); Kevin Hurler, *Here are All the Companies Using ChatGPT... So Far*, YAHOO! (May 24, 2023), [https://news.yahoo.com/companies-using-chatgpt-far-](https://news.yahoo.com/companies-using-chatgpt-far-205500883.html)
205500883.html.

¹² *Spotlight: AI Myths and Misconceptions—Transcript*, STENO (May 11, 2023),
<https://steno.ai/your-undivided-attention/spotlight-ai-myths-and-misconceptions>.

1 and we don't understand how it works. If we are not careful, we are setting ourselves up for a lot of
2 correlated failures.”¹³

3 15. Such aggressive deployment of Defendants' AI is reckless, without the proper
4 safeguards in place. “No matter how tall the skyscraper of benefits that AI assembles for us... if
5 those benefits land in a society that does not work anymore, because banks have been hacked, and
6 people's voices have been impersonated, and cyberattacks have happened everywhere and people
7 don't know what's true [... or] what to trust, [...] how many of those benefits can be realized in a
8 society that is *dysfunctional*?”¹⁴

9 16. Through their AI Products, integrated into every industry, Defendants collect, store,
10 track, share, and disclose **Private Information** of millions of users (“Users”), including: (1) all
11 details entered into the Products; (2) account information users enter when signing up; (3) name;
12 (4) contact details; (5) login credentials; (6) emails; (7) payment information for paid users; (8)
13 transaction records; (9) identifying data pulled from users' devices and browsers, like IP addresses
14 and location, including geolocation of the users; (10) social media information; (11) chat log data;
15 (12) usage data; (13) analytics; (14) cookies;¹⁵ (15) key strokes; and (16) typed searches, as well as
16 other online activity data. Defendants, through the Products, unlawfully obtain access to and
17 intercept this information from the individual users of applications and devices that have integrated
18 ChatGPT-4—including but not limited to user locations and image-related data obtained through
19 Snapchat,¹⁶ user financial information through Stripe, musical tastes and preferences through

20 _____
21 ¹³ *Id.*; see also Jan Leike (@janleike), TWITTER (May 17, 2023, 10:56 AM),
<https://twitter.com/janleike/status/1636788627735736321>.

22 ¹⁴ *Spotlight: AI Myths and Misconceptions—Transcript*, *supra* note 12.

23 ¹⁵ *Privacy Policy*, OPENAI, <https://openai.com/policies/privacy-policy> (last updated June 23, 2023).

24 ¹⁶ Jeremy Kahn & Kylie Robison, *Snap's 'My AI' Chatbot Tells Users it Doesn't Know Their*
Location. It Does, FORTUNE (Apr. 21, 2023), [https://fortune.com/2023/04/21/snap-chat-my-ai-lies-](https://fortune.com/2023/04/21/snap-chat-my-ai-lies-location-data-a-i-ethics/)
location-data-a-i-ethics/; *I Got Snapchat AI to Admit Everything*, REDDIT (May 20, 2023),
25 https://www.reddit.com/r/ChatGPT/comments/13gty7u/i_got_snapchat_ai_to_admit_everything/;
Snapchats New "My AI" Correctly Identifying Images it Claims it Can't View, Then Walks it
Back, REDDIT (Apr. 20, 2023),
26 [https://www.reddit.com/r/mildlyinfuriating/comments/12tdmzq/snapchats_new_my_ai_correctly_](https://www.reddit.com/r/mildlyinfuriating/comments/12tdmzq/snapchats_new_my_ai_correctly_identifying_images/)
identifying_images/; *Snapchat AI Can Determine What's In The Pictures You Send It*, REDDIT
27 (Apr. 20, 2023),
28 [https://www.reddit.com/r/oddlyterrifying/comments/12szymo/snapchat_ai_can_determine_whats_](https://www.reddit.com/r/oddlyterrifying/comments/12szymo/snapchat_ai_can_determine_whats_in_the_pictures/)
in_the_pictures/.

1 Spotify,¹⁷ user patterns and private conversation analysis through Slack and Microsoft Teams,¹⁸ and
 2 even private health information obtained through the management of patient portals such as
 3 MyChart.¹⁹

4 17. All of this personal information is captured in real time. Together with Defendants’
 5 scraping of our digital footprints—comments, conversations we had online yesterday, as well as 15
 6 years ago—Defendants now have enough information to create our digital clones, including the
 7 ability to replicate our voice and likeness and predict and manipulate our next move using the
 8 technology on which the Products were built. They can also misappropriate our skill sets and
 9 encourage our own professional obsolescence. This would obliterate privacy as we know it and
 10 highlights the importance of the privacy, property, and other legal rights this lawsuit seeks to
 11 vindicate.²⁰

12 18. Defendants must not only be enjoined from their ongoing violations of the privacy
 13 and property rights of millions, but they must also be required to take immediate action to implement
 14 proper safeguards and regulations for the Products, their users, and all of society, such as:

- 15 (i) **Transparency:** OpenAI should open the “black box,” to clearly and precisely disclose
 16 the data it is collecting, including where and from whom, in clear and conspicuous

17
 18 ¹⁷ Shlomo Sprung, *Spotify Introduces AI DJ Powered by ChatGPT Maker OpenAI*, BOARDROOM
 19 (Feb. 22, 2023), <https://boardroom.tv/spotify-ai-dj-chatgpt/> (ChatGPT in Spotify creates an “AI
 DJ” that utilizes Spotify’s algorithmic learnings to track users’ musical tastes and predict a
 personalized music lineup).

20 ¹⁸ Brad Lightcap, *How OpenAI Connects with Customers and Expands ChatGPT with Slack*,
 21 SLACK, [https://slack.com/customer-stories/openai-connects-with-customers-and-expands-chatgpt-
 with-slack](https://slack.com/customer-stories/openai-connects-with-customers-and-expands-chatgpt-with-slack) (last visited June 8, 2023); Ryan Morrison, *Microsoft to Integrate ChatGPT into
 Teams*, TECH MONITOR (May 4, 2023), [https://techmonitor.ai/technology/ai-and-
 automation/microsoft-to-integrate-chatgpt-into-teams](https://techmonitor.ai/technology/ai-and-automation/microsoft-to-integrate-chatgpt-into-teams) (explaining that ChatGPT will be able to
 22 automate notes and recommend tasks based on **verbal conversations** through Teams).

23 ¹⁹ Naomi Diaz, *6 Hospitals, Health Systems Testing out ChatGPT*, BECKER’S HEALTH IT (June 2,
 24 2023), [https://www.beckershospitalreview.com/innovation/4-hospitals-health-systems-testing-out-
 chatgpt.html](https://www.beckershospitalreview.com/innovation/4-hospitals-health-systems-testing-out-chatgpt.html).

25 ²⁰ Joanna Stern, *I Cloned Myself With AI. She Fooled My Bank and My Family*, WALL ST. J. (Apr.
 26 28, 2023, 7:58 AM), [https://www.wsj.com/articles/i-cloned-myself-with-ai-she-fooled-my-bank-
 and-my-family-356bd1a3](https://www.wsj.com/articles/i-cloned-myself-with-ai-she-fooled-my-bank-and-my-family-356bd1a3); Michael Atleson, *Chatbots, Deepfakes, and Voice Clones: AI
 Deception for Sale*, FED. TRADE COMM’N,(2023), [https://www.ftc.gov/business-
 guidance/blog/2023/03/chatbots-deepfakes-voice-clones-ai-deception-sale](https://www.ftc.gov/business-guidance/blog/2023/03/chatbots-deepfakes-voice-clones-ai-deception-sale); Dongwook Yoon, *AI
 Clones Made from User Data Pose Uncanny Risks*, THE CONVERSATION (June 4, 2023, 7:19 AM),
 27 <https://theconversation.com/ai-clones-made-from-user-data-pose-uncanny-risks-206357>.

1 policy documents that are explicit about how this information is to be stored, handled,
2 protected, and used;

3 (ii) **Accountability:** The developers of ChatGPT and the other AI Products should be
4 responsible for Product actions and outputs and barred from further commercial
5 deployment absent the Products' ability to follow a code of human-like ethical
6 principles and guidelines and respect for human values and rights, and until Plaintiffs
7 and Class Members are fairly compensated for the stolen data on which the Products
8 depend;

9 (iii) **Control:** Defendants must allow Product users and everyday internet users to opt out
10 of *all* data collection and they should otherwise stop the illegal taking of internet data,
11 delete (or compensate for) any ill-gotten data, or the algorithms which were built on
12 the stolen data, and before any further commercial deployment, technological safety
13 measures must be added to the Products that will prevent the technology from
14 surpassing human intelligence and harming others.

15 **PARTIES**

16 **Plaintiff A.T.**

17 19. Plaintiff A.T. is and at all relevant times was a resident of the State of California.

18 20. Plaintiff A.T. is a product engineer and began using ChatGPT-3.5 on or about
19 November 2022. He is a current user of ChatGPT-3.5 and ChatGPT-4.0. Plaintiff A.T. accesses the
20 Products from his personal computer, cellular device, and work computer.

21 21. Plaintiff A.T. engaged with a variety of websites and social media applications prior
22 to 2021. Plaintiff A.T. has used his accounts on those platforms to post content, react to and
23 comment on others' content, re-post other users' content, and to save and compile information in
24 line with his interests. He posted photos of himself, family members, friends, and other media to
25 these websites and social media applications.

26 22. For many years, Plaintiff A.T. has had a Spotify account which he frequently uses to
27 listen to music and create unique playlists. Plaintiff A.T. regularly views videos on YouTube, posted
28 content, and commented on other users' videos. Plaintiff A.T. has a variety of social media accounts

1 on Twitter, Reddit, TikTok, Snapchat, Yelp, LinkedIn, as well as Crunchbase, Webflow, and other
2 technology-focused sites. Plaintiff A.T. published many posts on these internet accounts,
3 accompanied by commentary.

4 23. Plaintiff A.T. has also founded or co-founded at least four companies, the details of
5 which are summarized on those respective websites.

6 24. Plaintiff A.T. has also posted online about his political views, as well as frequently
7 asked and answered technical questions using his professional knowledge on various websites.

8 25. Plaintiff A.T. is concerned that Defendants have taken his skills and expertise, as
9 reflected in his online contributions, and incorporated them into Products that could someday result
10 in professional obsolescence for software engineers like him.

11 26. Plaintiff A.T. reasonably expected that the information that he exchanged with these
12 websites prior to their introduction would not be intercepted by any third-party looking to compile
13 and use all his information and data for commercial purposes. Plaintiff A.T. did not consent to the
14 use of his private information by third parties in this manner. Notwithstanding, Defendants stole
15 Plaintiff A.T.'s personal data from across this wide swath of online applications and platforms to
16 train the Products.

17 **Plaintiff J.H.**

18 27. Plaintiff J.H. is and at all relevant times was a resident of the State of New York.

19 28. Plaintiff J.H. is a consumer and began using ChatGPT-3.5 on or about December
20 2022. He is a current user of ChatGPT-3.5 and ChatGPT-4.0. Plaintiff J.H. accesses the Products
21 from his personal computer and cellular device.

22 29. Plaintiff J.H. engaged with a variety of websites and social media applications prior
23 to 2021. Plaintiff J.H. has used his accounts on those platforms to post content, react to and comment
24 on others' content, re-post other users' content, and to save and compile information in line with his
25 interests. He had posted photos of himself, family members, friends, and other media to these
26 websites and social media applications, but removed them in December 2022 expecting that they
27 would remain in his exclusive control and not be used without his authorization.

28 30. For many years, Plaintiff J.H. has had a Spotify account which he frequently uses to

1 listen to podcasts. Plaintiff J.H. regularly views videos on YouTube, and has posted content to
2 YouTube. Plaintiff J.H. has a variety of social media accounts on Reddit, Twitter, TikTok,
3 Facebook, Snapchat, Instagram, Discord, and Yelp. Plaintiff J.H. published many posts on these
4 internet accounts, accompanied by commentary.

5 31. Plaintiff J.H. has also posted online about his personal beliefs, as well as seeking
6 advice and providing feedback with myriad users on various platforms.

7 32. Plaintiff J.H. is concerned that Defendants have taken his skills and expertise, as
8 reflected in his online contributions, and incorporated them into Products that could someday result
9 in professional obsolescence for software engineers like him.

10 33. Plaintiff J.H. reasonably expected that the information that he exchanged with these
11 websites prior to their introduction would not be intercepted by any third-party looking to compile
12 and use all his information and data for commercial purposes. Plaintiff J.H. did not consent to the
13 use of his private information by third parties in this manner. Notwithstanding, Defendants stole
14 Plaintiff J.H.'s personal data from across this wide swath of online applications and platforms to
15 train the Products.

16 **Defendants**

17 34. **Defendant OpenAI** is an AI research laboratory consisting of the non-profit OpenAI
18 Incorporated ("OpenAI Inc.") and its for-profit subsidiary corporation OpenAI Limited Partnership
19 ("OpenAI LP") (hereinafter, collectively, "OpenAI").²¹ OpenAI was founded in 2015 and is
20 headquartered in San Francisco, CA. OpenAI has released the AI-based products DALL-E, GPT-4,
21 OpenAI Five, ChatGPT, and OpenAI Codex for commercial (to integrate within one's business)
22 and personal use.

23 35. OpenAI was originally founded as a nonprofit research laboratory with a single
24 mission: "to advance [artificial] intelligence in the way that is most likely to benefit humanity as a
25 whole."²² In the words of OpenAI at the time, it was critical for the organization to be
26

27 _____
28 ²¹ *OpenAI LP*, OPENAI, <https://openai.com/blog/openai-lp> (last visited June 27, 2023).

²² Greg Brockman & Ilya Sutskever, *Introducing OpenAI*, OPENAI (Dec. 11, 2015),
<https://openai.com/blog/introducing-openai>.

1 “unconstrained by a need to generate a financial return.”²³ Fast forward to April of 2023: OpenAI
2 closed a more than \$300 million share sale at a valuation between \$27 billion and \$29 billion.²⁴
3 OpenAI projects that its AI chatbot, ChatGPT, will generate a revenue of \$200 million in 2023 and
4 exponentially grow to \$1 billion by the end of 2024.²⁵

5 36. Defendant OpenAI GP, L.L.C. (“OpenAI GP”) is a Delaware limited liability
6 company with its principal place of business located at 3180 18th Street, San Francisco, CA 94110.
7 OpenAI GP is wholly owned and controlled by OpenAI, Inc. Further, OpenAI GP is the general
8 partner of OpenAI, L.P. and is responsible for managing and operating the day-to-day business and
9 affairs of OpenAI, L.P. Its primary focus is research and technology. OpenAI GP was aware of the
10 unlawful conduct alleged herein and exercised control over OpenAI, L.P. throughout the Class
11 Period. OpenAI GP is liable for the debts, liabilities, and obligations of OpenAI, L.P., including
12 litigation and judgments.

13 37. Defendant OpenAI Startup Fund I, L.P. (“OpenAI Startup Fund I”) is a Delaware
14 limited partnership with its principal place of business located at 3180 18th Street, San Francisco,
15 CA 94110. Upon information and belief, OpenAI Startup Fund I played a vital role in the foundation
16 of OpenAI, L.P., including providing initial funding and creating its business strategy. By
17 participating in OpenAI Startup Fund I, certain entities and individuals obtained an ownership
18 interest in OpenAI, L.P. OpenAI Startup Fund I exercised control over OpenAI, L.P. and was aware
19 of the unlawful conduct alleged herein throughout the Class Period.

20 38. Defendant OpenAI Startup Fund GP I, L.L.C. (“OpenAI Startup Fund GP I”) is a
21 Delaware limited liability company with its principal place of business located at 3180 18th Street,
22 San Francisco, CA 94110. OpenAI Startup Fund GP I is the general partner of OpenAI Startup Fund
23 I and is responsible for managing and operating the day-to-day business and affairs of OpenAI
24 Startup Fund I. OpenAI Startup Fund GP I is liable for the debts, liabilities, and obligations of

25 _____
26 ²³ *Id.*

27 ²⁴ *OpenAI Closes \$300 Million Funding Round at \$27 Billion-\$29 Billion Valuation, TechCrunch*
reports, REUTERS (Apr. 28, 2023), <https://www.reuters.com/markets/deals/openai-closes-10-bln-funding-round-27-bln-29-bln-valuation-techcrunch-2023-04-28/>.

28 ²⁵ Jeffrey Dastin, *Exclusive: ChatGPT Owner OpenAI Projects \$1 Billion in Revenue by 2024*,
REUTERS (Dec. 15, 2022), <https://www.reuters.com/business/chatgpt-owner-openai-projects-1-billion-revenue-by-2024-sources-2022-12-15/>.

1 OpenAI Startup Fund I, including litigation and judgments. OpenAI Startup Fund GP I was aware
 2 of the unlawful conduct alleged herein and exercised control over OpenAI, L.P. throughout the
 3 Class Period. Sam Altman, co-founder, CEO, and Board member of OpenAI, Inc. is the Manager
 4 of OpenAI Startup Fund GP I.

5 39. Defendant OpenAI Startup Fund Management, LLC (“OpenAI Startup Fund
 6 Management”) is a Delaware limited liability company with its principal place of business located
 7 at 3180 18th Street, San Francisco, CA 94110. OpenAI Startup Fund Management exercised control
 8 over OpenAI, L.P. throughout the Class Period and thus, was aware of the unlawful conduct alleged
 9 herein.

10 40. **Defendant Microsoft Corporation** (“Microsoft”) is a Washington corporation with
 11 its principal place of business located at One Microsoft Way, Redmond, Washington 98052.
 12 Microsoft partnered with OpenAI in 2016 with the goal to “democratize Artificial Intelligence.” In
 13 July 2019, Microsoft invested \$1 billion in OpenAI LP at a \$20 billion valuation.²⁶ In 2020,
 14 Microsoft became the exclusive licensee of OpenAI’s GPT-3 language model—despite OpenAI’s
 15 continued claims that its products are meant to benefit “humanity” at large. In October 2022, news
 16 reports stated OpenAI was “in advanced talks to raise more funding from Microsoft” at that same
 17 \$20 billion valuation.²⁷ Then, in January of 2023, Microsoft confirmed its extended partnership with
 18 OpenAI by investing \$10 billion into ChatGPT.²⁸ Prior to this \$10 billion dollar investment,
 19 Microsoft had invested \$3 billion into OpenAI in previous years.²⁹

20
 21 ²⁶ Hasan Chowdhury, *Microsoft’s Investment into ChatGPT’s Creator May be the Smartest \$1*
 22 *Billion Ever Spent*, BUS. INSIDER (Jan. 6, 2023), [https://www.businessinsider.com/microsoft-](https://www.businessinsider.com/microsoft-openai-investment-the-smartest-1-billion-ever-spent-2023-1)
 23 [openai-investment-the-smartest-1-billion-ever-spent-2023-1](https://www.businessinsider.com/microsoft-openai-investment-the-smartest-1-billion-ever-spent-2023-1); Dina Bass, *Microsoft Invests \$10*
 24 *Billion in ChatGPT Maker OpenAI*, BLOOMBERG (Jan. 23, 2023),
 25 [https://www.bloomberg.com/news/articles/2023-01-23/microsoft-makes-multibillion-dollar-](https://www.bloomberg.com/news/articles/2023-01-23/microsoft-makes-multibillion-dollar-investment-in-openai#xj4y7vzkg)
 26 [investment-in-openai#xj4y7vzkg](https://www.bloomberg.com/news/articles/2023-01-23/microsoft-makes-multibillion-dollar-investment-in-openai#xj4y7vzkg).

27 Aaron Holmes et al., *OpenAI, Valued at Nearly \$20 Billion, in Advanced Talks with Microsoft*
 28 *for More Funding*, THE INFO. (Oct. 20, 2022), [https://www.theinformation.com/articles/openai-](https://www.theinformation.com/articles/openai-valued-at-nearly-20-billion-in-advanced-talks-with-microsoft-for-more-funding)
 29 [valued-at-nearly-20-billion-in-advanced-talks-with-microsoft-for-more-funding](https://www.theinformation.com/articles/openai-valued-at-nearly-20-billion-in-advanced-talks-with-microsoft-for-more-funding).

30 ²⁸ *Microsoft Confirms Its \$10 Billion Investment into ChatGPT, Changing How Microsoft*
 31 *Competes with Google, Apple and Other Tech Giants*, FORBES (Jan. 27, 2023),
 32 [https://www.forbes.com/sites/qai/2023/01/27/microsoft-confirms-its-10-billion-investment-into-](https://www.forbes.com/sites/qai/2023/01/27/microsoft-confirms-its-10-billion-investment-into-chatgpt-changing-how-microsoft-competes-with-google-apple-and-other-tech-giants/?sh=4eea29723624)
 33 [chatgpt-changing-how-microsoft-competes-with-google-apple-and-other-tech-](https://www.forbes.com/sites/qai/2023/01/27/microsoft-confirms-its-10-billion-investment-into-chatgpt-changing-how-microsoft-competes-with-google-apple-and-other-tech-giants/?sh=4eea29723624)
 34 [giants/?sh=4eea29723624](https://www.forbes.com/sites/qai/2023/01/27/microsoft-confirms-its-10-billion-investment-into-chatgpt-changing-how-microsoft-competes-with-google-apple-and-other-tech-giants/?sh=4eea29723624).

35 ²⁹ Cade Metz, *Microsoft to Invest \$10 Billion in OpenAI, the Creator of ChatGPT*, THE N.Y.
 36 *TIMES* (Jan. 23, 2023), [https://www.nytimes.com/2023/01/23/business/microsoft-chatgpt-artificial-](https://www.nytimes.com/2023/01/23/business/microsoft-chatgpt-artificial-intelligence.html)
 37 [intelligence.html](https://www.nytimes.com/2023/01/23/business/microsoft-chatgpt-artificial-intelligence.html).

1 41. Microsoft’s continued investments, as well as introduction of ChatGPT on its multiple
2 platforms (Bing, Microsoft Teams, etc.) underscore the depth of its partnership with OpenAI.
3 Through these investments, Microsoft gained exclusive access to the entire OpenAI codebase.³⁰
4 Furthermore, Microsoft Azure also acts as the exclusive cloud service of OpenAI.³¹

5 42. As OpenAI’s largest investor and largest service provider—specifically in connection
6 with the development of ChatGPT—Microsoft exerts considerable control over OpenAI. Analysts
7 estimate OpenAI will add between \$30 billion and \$40 billion to Microsoft’s top line.

8 43. **Agents and Co-Conspirators.** Defendants’ unlawful acts were authorized, ordered,
9 and performed by Defendants’ respective officers, agents, employees, and representatives, while
10 actively engaged in the management, direction, and control of Defendants’ businesses and affairs.
11 Defendants’ agents operated under explicit and apparent authority of their principals. Each
12 Defendant, and their subsidiaries, affiliates, and agents operated as a single unified entity.

13 **JURISDICTION AND VENUE**

14 44. This Court has subject matter jurisdiction over the federal claims in this action,
15 namely the Electronic Communications Privacy Act and the Computer Fraud and Abuse Act,
16 pursuant to 28 U.S.C. § 1331.

17 45. This Court also has subject matter jurisdiction over this action pursuant to the Class
18 Action Fairness Act, 28 U.S.C § 1332(d), because this is a class action in which the amount in
19 controversy exceeds \$5,000,000, exclusive of interest and costs. There are millions of class
20 members as defined below, and minimal diversity exists because a significant portion of class
21 members are citizens of a state different from the citizenship of at least one Defendant.

22 46. This Court also has supplemental jurisdiction over the state law claims in this action
23 pursuant to 28 U.S.C. § 1367 because the state law claims form part of the same case or controversy
24 as those that give rise to the federal claims.

25 _____
26 ³⁰ Mohit Pandey, *OpenAI, a Data Scavenging Company for Microsoft*, AIM (Mar. 24, 2023),
<https://analyticsindiamag.com/openai-a-data-scavenging-company-for-microsoft/>.

27 ³¹ *Microsoft Confirms Its \$10 Billion Investment Into ChatGPT, Changing How Microsoft*
Competes With Google, Apple And Other Tech Giants, FORBES (Jan. 27, 2023),
28 [https://www.forbes.com/sites/qai/2023/01/27/microsoft-confirms-its-10-billion-investment-into-
chatgpt-changing-how-microsoft-competes-with-google-apple-and-other-tech-
giants/?sh=4eea29723624](https://www.forbes.com/sites/qai/2023/01/27/microsoft-confirms-its-10-billion-investment-into-chatgpt-changing-how-microsoft-competes-with-google-apple-and-other-tech-giants/?sh=4eea29723624).

1 47. Pursuant to 28 U.S.C. § 1391, this Court is the proper venue for this action because a
 2 substantial part of the events, omissions, and acts giving rise to the claims herein occurred in this
 3 District: Defendant OpenAI is headquartered in this District, all Defendants gain significant revenue
 4 and profits from doing business in this District, consumers sign up for ChatGPT accounts and
 5 provide ChatGPT with their sensitive information in this District, Class Members affected by this
 6 data misuse reside in this District, and Defendants employ numerous people in this District—a
 7 number of whom work specifically on making the decisions regarding the data privacy and handling
 8 of consumers’ data that are challenged in this Action. Each Defendant has transacted business,
 9 maintained substantial contacts, and/or committed overt acts in furtherance of the illegal scheme
 10 and conspiracy throughout the United States, including in this District. Defendants’ conduct had the
 11 intended and foreseeable effect of causing injury to persons residing in, located in, or doing business
 12 throughout the United States, including in this District.

13 48. Defendants are subject to personal jurisdiction in California based upon sufficient
 14 minimum contacts which exist between Defendants and California. Defendants are authorized to do
 15 and are doing business in California, and Defendants advertise and solicit business in California.
 16 Defendants have purposefully availed themselves of the protections of California law and should
 17 reasonably expect to be hauled into court in California for harm arising out of their pervasive
 18 contacts with the State. Further, for Defendant OpenAI, the decisions affecting consumers data and
 19 privacy stem from the company’s San Francisco office headquarters.

FACTUAL BACKGROUND

I. DEVELOPMENT OF ARTIFICIAL INTELLIGENCE IN THE U.S.

A. OpenAI: From Open Nonprofit to Profit-Driven \$29B Commercial Partner of Tech Giant Microsoft

24 49. OpenAI was founded in 2015 as a nonprofit research laboratory with a single mission:
 25 “to advance artificial intelligence in a way that would benefit society as a whole. . . .”³² Critical to
 26 that mission, according to OpenAI at the time, was for the organization to be “unconstrained by a
 27

28 ³² *The Transformation of OpenAI From Nonprofit to \$29B For-Profit*, THE SOCIABLE (Apr. 5, 2023), <https://sociable.co/business/the-transformation-of-openai-from-nonprofit-to-29b-for-profit/>.

1 need to generate a financial return.”³³ The nonprofit was thus funded by million-dollar donations
2 from prominent, wealthy entrepreneurs and researchers who shared the non-profit’s vision of
3 creating safe, ethical, and responsible AI, to benefit humankind and to do no harm, and who
4 recognized the dangers that could befall society if AI were developed and launched for commercial
5 gain.

6 50. OpenAI also originally pledged to “freely collaborate” with other responsible
7 organizations and researchers, in part by making its research available to inspect and audit as a
8 further “check” on the safety of any AI capabilities, to help ensure the powerful technology on
9 which they were working would not someday destroy lives and ultimately, civilization. The
10 founders believed this openness was so critical to the non-profit’s mission, that they named it
11 “Open” AI. As they further explained at the time, “since our research is free from financial
12 obligations, we can better focus on a positive human impact. We believe AI should be an extension
13 of individual human wills, and in the spirit of liberty, as broadly and evenly distributed as
14 possible.”³⁴

15 51. For years, OpenAI purported to operate as such: openly and in pursuit of its single
16 mission to advance humanity, safely and responsibly. That all changed in 2019, when OpenAI
17 abruptly “shut its doors” to all ‘Open’ influence and scrutiny, shifted to a profit-generating corporate
18 structure, and decided instead to focus on commercializing the AI capabilities on which it had been
19 working.

20 52. At the time, Google Brain’s “transformer” innovation had opened a new frontier in
21 AI development, where AI could improve endlessly, some experts believe to even superhuman
22 intelligence—but only if it were fed “endless data” to train it, a costly endeavor given the computing
23 power required.³⁵ To do so, OpenAI entered an exclusive partnership with Microsoft, which
24 invested \$1B into the company, gaining the only outside access to the effort once “Open” to all.

25
26 ³³ *Id.*

27 ³⁴ Greg Brockman & Ilya Sutskever, *Introducing OpenAI*, OPENAI (Dec. 11, 2015),
<https://openai.com/blog/introducing-openai>.

28 ³⁵ Reed Albergotti, *The Secret history of Elon Musk, Sam Altman, and OpenAI*, SEMAFOR (Mar.
24, 2023), <https://www.semafor.com/article/03/24/2023/the-secret-history-of-elon-musk-sam-altman-and-openai>.

1 Together, they built a “supercomputer” to train massive language models that ultimately resulted in
2 ChatGPT and the image generator DALL-E.³⁶

3 53. OpenAI’s sudden shift to a profit focus and alignment with Microsoft, a corporate
4 giant with a vested interest in curating and dominating a commercial market for AI, marked the
5 beginning of the end of OpenAI’s commitment to humanity. The company began to pursue profits
6 at the expense of privacy, security, and ethics, beginning with its data collection.

7 54. To realize the most powerful and thus most profitable AI, OpenAI would need data,
8 and lots of it, to “train” the language models on which the Products run using the supercomputer it
9 had built in partnership with Microsoft. Defendants thus doubled down on their strategy to secretly
10 harvest millions of consumers’ personal data from the internet. Then, on the backs of this stolen
11 data, they rushed to market the Products without adequate safeguards or controls to ensure their
12 safety. While Defendants recognized then, as they do now, that they cannot fully predict how the
13 Products might evolve to operate, they knew the public would be amazed by the Products already
14 seemingly near human “intelligence” and other capabilities. And thus, they knew they could make
15 a ton of money.

16 55. In public, OpenAI continued to state its commitment to ethical AI development. But
17 with its new profit orientation, that “was kind of like trying to juggle while riding a unicycle, except
18 with more existential questions about the nature of humanity.”³⁷ Defendants acknowledge they do
19 not understand the full scope of the risks posed by the Products currently, and no one knows how
20 AI might evolve now that billions of people are using the technology every day.³⁸ Defendants, like

21 _____
22 ³⁶ *Id.*

23 ³⁷ *The Transformation of OpenAI From Nonprofit to \$29B For-Profit*, THE SOCIABLE (Apr. 5,
24 2023), <https://sociable.co/business/the-transformation-of-openai-from-nonprofit-to-29b-for-profit/>.

25 ³⁸ “As a system like this learns from data, it develops skills that its creators never expected. It is
26 hard to know how things might go wrong after millions of people start using it.” *See* Cade Metz,
27 *What’s the Future for AI?*, THE N.Y. TIMES (Mar. 31, 2023),
28 <https://www.nytimes.com/2023/03/31/technology/ai-chatbots-benefits-dangers.html>; Jason
Abbruzzese, *The Tech Watchdog that Raised Alarms About Social Media is Warning About AI*,
NBC NEWS (Mar. 22, 2023), <https://www.nbcnews.com/tech/tech-news/tech-watchdog-raised-alarms-social-media-warning-ai-rcna76167> (“What’s surprising and what nobody foresaw is that
just by learning to predict the next piece of text on the internet, these models are developing new

1 other leading experts, are united in believing the ultimate risk posed by AI is the collapse of
 2 civilization as we know it. And yet, they released the Products worldwide anyway, setting off a
 3 global AI arms race.

4 56. Earlier this year, OpenAI raised another \$10B from its single corporate partner,
 5 Microsoft, increasing its then corporate valuation to \$29B and giving Microsoft a significant stake
 6 in the company. With that, the 180-degree transformation—from open nonprofit for the benefit of
 7 humanity to closed corporate profit machine fueled by greed and market power—was complete.

8 57. OpenAI’s shift in organizational structure has raised eyebrows given its
 9 unprecedented nature, and the moral and legal questions it raises. AI researchers, ethicists, and the
 10 public share concerns about the conflict between OpenAI’s original mission to benefit humanity on
 11 the one hand and the current profit-driven motives of investors, chiefly Microsoft, on the other.³⁹
 12 They worry that OpenAI is prioritizing short-term financial gains over long-term safety and ethical
 13 considerations, as exemplified by the sudden deployment of the Products for widespread
 14 commercial use despite all the known dangers.⁴⁰ Moreover, as one commentator noted, “there are
 15 various different ways to make hundreds of millions of dollars, but historically ‘starting a nonprofit’
 16 has not been one of them.”⁴¹

17 58. Elon Musk, an original non-profit funder and founder, was more blunt as to the
 18 seismic shift: “I’m still confused as to how a non-profit to which I donated ~100M somehow became
 19 a \$30B market cap for-profit.” He noted, “OpenAI was created as an open source (which is why I
 20 named it ‘Open’ AI), non-profit company to serve as a counterweight to Google, but now it has

21 _____
 22 capabilities that no one expected. . . So just by learning to predict the next character on the
 23 internet, it’s learned how to play chess.” Others have also commented on the technology
 24 continuing to display unintended and unpredictable emergent capabilities. Jason Wei, *137*
 25 *Emergent Abilities of Large Language Models*, JASON WEI (Nov. 14, 2022),
 26 <https://www.jasonwei.net/blog/emergence>; Stephen Ornes, *The Unpredictable Abilities Emerging*
 27 *from Large AI Models*, QUANTA MAG. (Mar. 16, 2023), [https://www.quantamagazine.org/the-](https://www.quantamagazine.org/the-unpredictable-abilities-emerging-from-large-ai-models-20230316/)
 28 [unpredictable-abilities-emerging-from-large-ai-models-20230316/](https://www.quantamagazine.org/the-unpredictable-abilities-emerging-from-large-ai-models-20230316/).

³⁹ *From Non-Profit to Profit Monster: OpenAI’s Controversial Corporate Shift*, EXPLORING
 CHATGPT (Apr. 8, 2023), [https://exploringchatgpt.substack.com/p/from-non-profit-to-profit-](https://exploringchatgpt.substack.com/p/from-non-profit-to-profit-monster)
 monster.

⁴⁰ *Id.*

⁴¹ Felix Salmon, *How a Silicon Valley Nonprofit Became Worth Billions*, AXIOS (Jan. 10, 2023),
<https://www.axios.com/2023/01/10/how-a-silicon-valley-nonprofit-became-worth-billions>.

1 become a closed source, maximum profit company effectively controlled by Microsoft.”⁴²

2 59. If soliciting non-profit contributions to then turn around and build a for-profit
3 company “is legal,” Musk opined, then “why doesn’t everyone do it?”⁴³ This same question must
4 be asked about the equally unprecedented theft of personal data that is at the heart of this Action,
5 and the answer to both questions is the same: *It isn’t*.

6 60. As explained below, the only thing still ‘open’ about OpenAI is its open disregard for
7 the privacy and property interests of hundreds of millions. Worse, as a result of OpenAI’s
8 machinations for profit, “the most powerful tool mankind has ever created, is now in the hands of a
9 ruthless corporate monopoly.”⁴⁴

10 **B. OpenAI’s Products**

11 61. The most well-known of OpenAI’s products—and of all AI worldwide—is the
12 ground-breaking chatbot, ChatGPT. Once users input a question or a prompt in ChatGPT, the
13 information is digested by the AI model and the chatbot produces a response based on the
14 information a user has given and how that fits into its vast amount of training data.

15 62. ChatGPT was released as a “research preview” on November 30, 2022.⁴⁵ A blog post
16 casually introduced the AI chatbot to the world, thusly: “We’ve trained a model . . . which interacts
17 in a conversational way.” ChatGPT subsequently exploded in popularity, reaching **100 million**
18 **users** in only two months, making it the fastest-growing app in history.⁴⁶ For comparison, TikTok
19 took nine months to reach the same benchmark.⁴⁷ ChatGPT has continued to evolve exponentially,
20

21 ⁴² Sawdah Bhaimiya, *OpenAI Cofounder Elon Musk Said the Non-Profit He Helped Create is Now*
22 *Focused on ‘Maximum-Profit,’ Which is ‘Not What I Intended at All’*, BUS. INSIDER (Feb. 17,
23 2023), https://www.businessinsider.com/elon-musk-defends-role-in-openai-ChatGPT-microsoft-2023-2?utm_source=flipboard&utm_content=user%2FInsiderBusiness.

24 ⁴³ @elonmusk, TWITTER (Mar. 15, 2023),

<https://twitter.com/elonmusk/status/1636047019893481474>.

25 ⁴⁴ Marvie Basilan, *Elon Musk Says He’s The Reason OpenAI Exists as Sam Altman Testifies*
26 *Before Congress*, INT’L BUS. TIMES (May 17, 2023), <https://www.ibtimes.com/elon-musk-says-hes-reason-openai-exists-sam-altman-testifies-before-congress-3693771>.

27 ⁴⁵ *Introducing ChatGPT*, OPENAI (NOV. 30, 2022), <https://openai.com/blog/chatgpt>.

28 ⁴⁶ Krystal Hu, *ChatGPT Sets Record for Fastest-Growing User Base - Analyst Note*, REUTERS (Feb. 2, 2023), <https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/>.

⁴⁷ *Id.*

1 **with 1.8 billion visits in April of 2023.**⁴⁸

2 63. ChatGPT was built on a family of large language models (“LLMs”) collectively
3 known as GPT-3. As explained below, ChatGPT-3.5 was trained on 570GB of text data from the
4 internet containing hundreds of billions of words,⁴⁹ including text harvested from books, articles,
5 and websites, including social media. Due to its vast training data, ChatGPT can generate human-
6 like answers to text prompts and questions making it interact like “a friendly robot.”⁵⁰ On command
7 it can do a lot of what people do, like write poetry, compose music, draft research papers, create
8 lesson plans, and so much more, only faster than one human ever could. Naturally, the world was
9 stunned by these capabilities.

10 64. OpenAI has also released other AI-based products DALL-E, OpenAI Five, and
11 OpenAI Codex for commercial (to integrate within one’s business) and personal use. It also
12 developed a program VALL-E, which has not been released for use to the public yet.

13 65. DALL-E (consisting of DALL-E and DALL-E 2) are deep learning models developed
14 by OpenAI to generate realistic digital images from natural language descriptions, known as
15 “prompts.”⁵¹ DALL-E uses a version of GPT-3, modified to generate images.⁵²

16 66. OpenAI Five is a computer program developed by OpenAI that plays the five-on-five
17 video game Dota 2.⁵³

18 67. OpenAI Codex is another artificial intelligence model developed by Open AI, which
19 is programmed to generate computer code for use in programming applications.⁵⁴

20 68. VALL-E is another artificial intelligence model intended to synthesize high-quality

21 ⁴⁸ Nerdynav, *97+ ChatGPT Statistics & User Numbers in June 2023 (New Data)*, NERDY NAV
22 (June 2, 2023), <https://nerdynav.com/chatgpt-statistics/>.

23 ⁴⁹ Uri Gal, *CHATGPT Collected Our Data Without Permission and is Going to Make Billions Off it*, SCROLL.IN (Feb. 15, 2023), <https://scroll.in/article/1043525/chatgpt-collected-our-data-without-permission-and-is-going-to-make-billions-off-it>.

24 ⁵⁰ Mark Wilson, *ChatGPT Explained: Everything You Need to Know About the AI Chatbot*,
25 TECHRADAR (Mar. 15, 2023), <https://www.techradar.com/news/chatgpt-explained>.

26 ⁵¹ Khari Johnson, *OpenAI Debuts DALL-E for Generating Images from Text*, VENTURE BEAT (Jan.
27 5, 2021), <https://venturebeat.com/business/openai-debuts-dall-e-for-generating-images-from-text/>.

28 ⁵² *Id.*

⁵³ Ben Dickson, *AI Defeated Human Champions at Dota 2*, TECHTALKS (Apr. 17, 2019),
<https://bdtechtalks.com/2019/04/17/openai-five-neural-networks-dota-2/>.

⁵⁴ Thomas Smith, *Why OpenAIs Codex Won’t Replace Coders*, IEEE SPECTRUM (Sept. 28, 2021),
<https://spectrum.ieee.org/openai-wont-replace-coders>.

1 personalized speech utilizing only a 3-second enrolled recording of an unseen speaker as a prompt.⁵⁵
 2 VALL-E was trained on audio voices from thousands of speakers.⁵⁶

3 **C. ChatGPT’s Development Depends on Secret Web-Scraping**

4 69. The large language models responsible for the Products depend on consuming huge
 5 amounts of data, in order to “train” the AI. Valuable to the process is personal data of any kind,
 6 including conversational data between humans, as this is how the Products develop what appear to
 7 be such human-like capabilities.

8 70. As a general matter, internet user data is available for purchase like any other content
 9 or property. In the technological era in which we live, a mature market for such data exists given
 10 how valuable our personal information has become to companies, for marketing and other purposes.
 11 The legal acquisition of data typically depends on consent and remuneration, with some form of
 12 consideration exchanged.

13 71. Despite established protocols for the purchase and use of personal information,
 14 Defendants took a different approach: *theft*. They systematically scraped 300 billion words from the
 15 internet, “books, articles, websites and posts – including personal information obtained without
 16 consent.”⁵⁷ OpenAI did so in secret, and without registering as a data broker as it was required to
 17 do under applicable law (*See infra* at Section III.A).

18 72. “Scraping involves the use of ‘bots,’ or robot applications deployed for automated
 19 tasks, which scan and copy the information on webpages then *store* and *index* the information.”⁵⁸
 20 According to a computer science professor at the University of Oxford, Michael Wooldridge, the
 21

22
 23 ⁵⁵ VALL-E Neural Codec Language Models are Zero-Shot Text to Speech Synthesizers, GITHUB
 PAGES, <https://lifeiteng.github.io/vall-e/index.html> (last visited June 27, 2023).

24 ⁵⁶ VALL-E: Five Things to Know About Microsoft’s AI Model That Can Mimic Any Voice in Three
 25 Seconds, TIMES OF INDIA (Jan. 11, 2023), <https://timesofindia.indiatimes.com/gadgets-news/vall-e-5-things-to-know-about-microsofts-ai-model-that-can-mimic-any-voice-in-3-seconds/articleshow/96898774.cms>.

26 ⁵⁷ Uri Gal, *ChatGPT is a Data Privacy Nightmare. If You’ve Ever Posted Online, You Ought to be*
 27 *Concerned*, THE CONVERSATION (Feb. 7, 2023), <https://theconversation.com/chatgpt-is-a-data-privacy-nightmare-if-youve-ever-posted-online-you-ought-to-be-concerned-199283>.

28 ⁵⁸ Will Hillier, *What is Web Scraping? A Complete Beginners Guide*, CAREER FOUNDRY (Aug. 13,
 2021), <https://careerfoundry.com/en/blog/data-analytics/web-scraping-guide/>.

1 full extent of personal data taken by Defendants’ scraping is “unimaginable.”⁵⁹

2 73. In his interview with The Guardian, Professor Wooldridge explained that the LLM
3 underlying ChatGPT, and other AIs like it, “includes the whole of the world wide web – *everything*.
4 Every link is followed in every page, and every link in those pages is followed.”⁶⁰ Thus, swept up
5 into the Products is “a lot of data about you and me.”⁶¹ Others have noted that the data includes
6 transcripts of our online chat logs, from across the internet, and other forms of personal conversation
7 such as our online customer service interactions and social media conversations, as well as “billions
8 of images scraped from the internet.”⁶² Many of these images were of “children and came from
9 photo sites and personal blogs.”⁶³

10 74. The unprecedented scope of the effort together with Defendants’ failure to seek
11 consent has been described as “the elephant in the room. . . all this training data must come from
12 somewhere. ChatGPT has effectively scraped the entire internet[.]”⁶⁴ As a result, Defendants have
13 essentially embedded into the Products personal information across a range of categories that reflect
14 our hobbies and interests, our religious beliefs, our political views and voting records, the social and
15 support groups to which we belong, our sexual orientations and gender identities, our personal
16 relationship statuses, our work information and histories, details (including pictures) about our
17 families and children, the music we listen to, our purchasing behaviors, our general likes and
18 dislikes, the ways in which we speak and write, our mental health and ailments, where we live and
19 where we go, the websites we visit, our digital subscriptions, our friend groups and other
20

21 ⁵⁹ Alex Hern & Dan Milmo, *I Didn’t Give Permission: Do AI’s Backers Care About Data Law*
22 *Breaches?*, THE GUARDIAN (Apr. 10, 2023),
23 [https://www.theguardian.com/technology/2023/apr/10/i-didnt-give-permission-do-ais-backers-](https://www.theguardian.com/technology/2023/apr/10/i-didnt-give-permission-do-ais-backers-care-about-data-law-breaches)
24 [care-about-data-law-breaches](https://www.theguardian.com/technology/2023/apr/10/i-didnt-give-permission-do-ais-backers-care-about-data-law-breaches).

⁶⁰ *Id.*

⁶¹ *Id.*

⁶² Jit Roy, *Data Source of ChatGPT*, ABOUTCHATGPT.COM (Jan. 2, 2023),
25 <https://aboutchatgpt.com/data-source-of-chatgpt/>; *see also* Hern & Milmo, *supra* note 59.

⁶³ Drew Harwell, *AI-generated child sex images spawn new nightmare for the web*, THE WASH.
26 POST (June 19, 2023), [https://www.msn.com/en-us/news/us/ai-generated-child-sex-images-spawn-](https://www.msn.com/en-us/news/us/ai-generated-child-sex-images-spawn-new-nightmare-for-the-web/ar-AA1cKhLH)
27 [new-nightmare-for-the-web/ar-AA1cKhLH](https://www.msn.com/en-us/news/us/ai-generated-child-sex-images-spawn-new-nightmare-for-the-web/ar-AA1cKhLH).

⁶⁴ Deep Tech Insights, *ChatGPT is a Threat, but Google is Still a Buy*, SEEKING ALPHA (Dec. 19,
28 2022), [https://seekingalpha.com/article/4565302-alphabet-ChatGPT-is-a-threat-but-google-is-still-](https://seekingalpha.com/article/4565302-alphabet-ChatGPT-is-a-threat-but-google-is-still-a-buy)
[a-buy](https://seekingalpha.com/article/4565302-alphabet-ChatGPT-is-a-threat-but-google-is-still-a-buy).

1 associational data, our email addresses, other contact and identifying information, and more.⁶⁵ With
 2 respect to personally identifiable information, Defendants fail sufficiently to filter it out of the
 3 training models, putting millions at risk of having that information disclosed on prompt or otherwise
 4 to strangers around the world.⁶⁶

5 75. The breadth and scope of Defendants’ data collection without permission, impacting
 6 essentially every internet user ever, raises serious legal, moral, and ethical issues.⁶⁷ One critique
 7 summarized the privacy risk bluntly, as follows: “*ChatGPT is a data privacy nightmare. If you’ve*
 8 *ever posted online, you ought to be concerned.*”⁶⁸ While regulators and courts around the world
 9 seek to crack down on AI researchers “hoovering up content without consent or notice,” the
 10 response, by Defendants and others, has been to keep their datasets largely secret, and to not grant
 11 regulator or other audit access.⁶⁹

12 76. Despite “*Open*” AI’s “absolute secrecy” surrounding its data collections and
 13 practices,⁷⁰ we know at the highest levels that the Company used (at least) five (5) distinct datasets
 14 to train ChatGPT: (1) Common Crawl; (2) WebTex2, text of webpages from all outbound Reddit
 15

16 ⁶⁵ *Digital Footprint: What is It And Why You Should Care About It*, INVISIBLY (Jan. 25, 2022),
 17 <https://www.invisibly.com/learn-blog/digital-footprint/> (“Your digital footprint is your trail of
 18 personal information that companies can follow. . . . To break it down, your digital footprint is
 19 essentially a record of your online activity. Whenever you log into an account, send an email, or
 20 buy something online, it leaves a digital impression behind. It is the trail of data left behind by
 21 your daily interactions. Your footprint is permanent which can leave your information vulnerable
 22 if not protected correctly. You might not always be aware that you are creating your digital
 footprint. For instance, websites can track your activity by installing cookies on your device.
 Furthermore, apps can collect your data without you even knowing it. Once an organization has
 access to your data, they can sell or share it with third parties. Even more, your information is out
 there and could be compromised via a data breach.”).

23 ⁶⁶ Katyanna Quach, *What happens when your massive text-generating neural net starts spitting out*
people’s phone numbers? If you’re OpenAI, you create a filter, THE REG. (Mar. 18, 2021),
 24 https://www.theregister.com/2021/03/18/openai_gpt3_data/?td=readmore-top.

25 ⁶⁷ Erin Griffith & Cade Metz, *A New Era of A.I. Booms, Even Amid the Tech Gloom*, THE N.Y.
 TIMES (Jan. 7, 2023), [https://www.nytimes.com/2023/01/07/technology/generative-ai-chatgpt-](https://www.nytimes.com/2023/01/07/technology/generative-ai-chatgpt-investments.html)
 26 [investments.html](https://www.nytimes.com/2023/01/07/technology/generative-ai-chatgpt-investments.html) (“The technology has raised thorny ethical questions around how generative A.I.
 may affect copyrights and whether the companies need to get permission to use the data that trains
 their algorithms.”).

27 ⁶⁸ Gal, *supra* note 57.

28 ⁶⁹ Hern & Milmo, *supra* note 59.

⁷⁰ *Id.* (“Copyright lawsuits and regulator actions against OpenAI are hampered by the company’s
 absolute secrecy about its training data.”).

1 links from posts with 3+ upvotes; (3) Books1; (4) Books2; and (5) Wikipedia.⁷¹

2 77. Of these training datasets, WebTex2 is OpenAI’s “proprietary” AI corpus of personal
3 data. To build it, OpenAI scraped every webpage linked to on the social media site Reddit in all
4 posts that received at least 3 “likes” (known as “Karma” votes on Reddit), together with the Reddit
5 posts and rich conversational data from its users around the world. The most popular “outbound”
6 links on Reddit include many of the most popular websites in the world, where people post personal
7 information, video, and audio clips of themselves and more, e.g., YouTube, Facebook, TikTok,
8 Snapchat, and Instagram. Given Defendants’ scraping protocols, all of this “outbound” data from
9 these various websites was targeted for taking, without notice or consent, to feed the large language
10 models on which the Products depend.

11 78. The co-founder and CEO of Reddit, Steve Huffman, remarked on the breadth of
12 Defendants’ unauthorized scraping, noting that he found it unacceptable that OpenAI has been
13 scraping “huge amounts of Reddit data to train their systems – for free.”⁷² According to Huffman,
14 “The Reddit corpus of data is really valuable. But we don’t need to give all of that value to some of
15 the largest companies in the world for free.”⁷³

16 79. Defendants’ theft related to their WebTex2 corpus is ongoing and continuous. As one
17 article explains, “the advantage of using the Webtext dataset is that it is constantly updated with
18 new data. As new web pages are added to the internet, they are included in the dataset, which helps
19 to ensure that the model is trained on the most recent and relevant language data.”⁷⁴ Neither Reddit
20 itself nor Reddit users, much less all the owners of the webpages and personal data linked to and
21 from Reddit, consent to this taking of data.

22 80. The other primary data set on which the Products depend, that the public currently
23 knows about, is the “Common Crawl,” a massive collection of web pages and websites also derived

24 ⁷¹ Patrick Meyer, *ChatGPT: How Does It Work Internally*, MEDIUM (Dec. 10, 2022),
25 <https://pub.towardsai.net/chatgpt-how-does-it-work-internally-e0b3e23601a1?gi=f28c10d5afef>.

26 ⁷² Gintaras Raauskas, *Redditors on Strike but Company Wants OpenAI to Pay Up for Scraping*,
27 CYBERNEWS, <https://cybernews.com/news/reddit-strike-api-openai-scraping/> (last updated June
12, 2023).

28 ⁷³ *Id.*

⁷⁴ GPTBlogs, *ChatGPT: How Much Data is Used in the Training Process?*, (Feb. 9, 2023),
<https://gptblogs.com/chatgpt-how-much-data-is-used-in-the-training-process>.

1 from large-scale web scraping. It contains petabytes of data collected over twelve (12) years,
2 including raw webpage data, metadata extracts, and text extracts from all types of websites.⁷⁵ In
3 total, the Common Crawl dataset constitutes nearly a trillion words.

4 81. The Common Crawl dataset is owned by a non-profit of the same name, which has
5 been indexing and storing as much of the World Wide Web as it can access, filing away as many as
6 3 billion webpages every month, for over a decade.⁷⁶ The non-profit makes the data available to the
7 public for free—but for research and educational purposes. As a result, the Common Crawl is a
8 staple of large *academic* studies of the web.⁷⁷ It was never intended to be taken *en masse* and turned
9 into an AI product for commercial gain, as Defendants have done. On information and belief, the
10 501(c)(3) overseeing the Common Crawl did not consent to this mass misappropriation of personal
11 data for commercial purposes. And even if it did, it did not obtain consent from internet users whose
12 personal data it scraped.

13 82. The commercial misappropriation of the Common Crawl has raised concerns given
14 the amount of personal data it contains, including highly personal data. One chilling example of the
15 privacy invasions caused by Defendants’ misappropriation is the experience of a San Francisco-
16 based digital artist named Lapine. Using the online tool “Have I Been Trained,” Lapine was able to
17 determine that her private medical file—i.e., photographs taken of her body as part of clinical
18 documentation when she was undergoing treatment for a rare genetic condition—ended up online
19 and then, memorialized in the Common Crawl archive.⁷⁸

20 83. Remarking on the web scraping practices in which Defendants engaged and the
21 subsequent commercialization of the ill-gotten data, Lapine highlighted the unique scope of the
22 harm: “It’s the digital equivalent of receiving stolen property. . . [my medical information] was
23

24 ⁷⁵ *Want to Use Our Data*, COMMON CRAWL, <https://commoncrawl.org/the-data/> (last visited June
25 27, 2023).

26 ⁷⁶ James Bridle, *The Stupidity of AI*, THE GUARDIAN (Mar. 16, 2023),
[https://www.theguardian.com/technology/2023/mar/16/the-stupidity-of-ai-artificial-intelligence-
dall-e-chatgpt](https://www.theguardian.com/technology/2023/mar/16/the-stupidity-of-ai-artificial-intelligence-dall-e-chatgpt).

27 ⁷⁷ Kalev Leetaru, *Common Crawl and Unlocking Web Archives for Research*, FORBES (Sept. 28,
28 2017), [https://www.forbes.com/sites/kalevleetaru/2017/09/28/common-crawl-and-unlocking-web-
archives-for-research/?sh=19e3c5373b83](https://www.forbes.com/sites/kalevleetaru/2017/09/28/common-crawl-and-unlocking-web-archives-for-research/?sh=19e3c5373b83).

⁷⁸ Bridle, *supra* note 76.

1 scraped into this dataset. . . it’s bad enough to have a photo leaked, *but now it’s part of a product.*”⁷⁹
 2 More broadly, this “productization” of personal information means all this data about us, scraped
 3 without permission, can now fuel ChatGPT’s responses to strangers around the world.⁸⁰ Worse,
 4 ChatGPT is the “new favorite toy” of online criminals, as the billions of personal and other data
 5 points about us, “scraped by ChatGPT, are now *free to use* for any number of targeted attacks,
 6 including malware, ransomware, phishing, Business Email Compromise, and social engineering.”⁸¹

7 84. As described further in Section III, this secret and unregistered scraping of internet
 8 data, for Defendants’ own private and exorbitant financial gain, without regard to privacy risks,
 9 amounts to the negligent and otherwise illegal theft of personal data of millions of Americans who
 10 do not even use AI tools. These individuals (“Non-Users”) had their personal information scraped
 11 long before OpenAI’s applications were available to the public, and certainly before they could have
 12 registered as a ChatGPT user. In either case, no one consented to the use of their personal data to
 13 train the Products.

14 85. OpenAI is now worth around \$29B, yet the individuals and companies that produced
 15 the data it scraped from the internet have not been compensated.⁸² This Action seeks to change that,
 16 and in the process, protect the privacy rights of millions.

17 **D. ChatGPT Training on Users of Defendants’ Programs and Applications.**

18 86. After using personal data taken without consent from millions of consumers to train
 19 the Products initially, Defendants continued to train the AI on data gleaned from ChatGPT’s
 20 registered users and users of ChatGPT plug-ins with sponsoring applications (“Users”). Defendants
 21 fed their AI models all of the data derived from User interactions—every click, entry, question, use,
 22 every move, key stroke, search, User’s geolocation (despite Users’ unwillingness to share that

23 ⁷⁹ *Id.*

24 ⁸⁰ *Is ChatGPT a Disaster for Data Privacy?*, BUS. REP. (Feb. 17, 2023), <https://www.business-reporter.co.uk/risk-management/is-chatgpt-a-disaster-for-data-privacy>.

25 ⁸¹ *Id.*

26 ⁸² Chris Morris, *OpenAI is Reportedly Raising Funds at a \$29 Billion Valuation—and its*
 27 *ChatGPT Could Challenge Google Search by Getting Wrapped into Microsoft Bing*, FORTUNE
 (Jan. 6, 2023), <https://fortune.com/2023/01/06/openai-valuation-ai-chatgpt-microsoft-bing-google-search/>;
 Jagmeet Singh & Ingrid Lunden, *OpenAI Closes \$300M Share Sale at \$27-29B Valuation*,
 28 TECH CRUNCH (Apr. 28, 2023), <https://techcrunch.com/2023/04/28/openai-funding-valuation-chatgpt/?tpcc=tcplustwitter>.

1 information)—as training data. Until recently, this also included all user interactions across the
2 hundreds or thousands of different platforms that now have ChatGPT plug-ins.

3 87. Following widespread criticism from consumers, OpenAI allegedly curtailed this
4 model of training their AI systems with user input, with CEO Sam Altman proclaiming broadly,
5 ***“Customers clearly want us not to train on their data, so we’ve changed our plans: We will not***
6 ***do that.”***⁸³ However, what OpenAI did not make clear is that, according to the updated Terms of
7 Use, it will only purportedly refrain from training on data from *API users*, but “[it] may use Content
8 from Services other than our API (“Non-API Content”) to help develop and improve our
9 Services.”⁸⁴ That means Defendants continue to feed the inputted, collected, and stored data of the
10 millions of everyday ChatGPT users to train the AI Products, despite the Company’s broad,
11 deliberately vague, and misleading pronouncement to the public that they “will not do that.” OpenAI
12 has also failed sufficiently to disclose that training aside (and even as to API users) it monitors,
13 saves, and shares all the personal information collected with its partners, including Microsoft.

14 88. ChatGPT’s systematic and intentional campaign to collect vast amounts of personal
15 information from Users without their knowledge or consent includes any information a user inputs
16 into the chat box with ChatGPT, as well as that user’s account information, contact details, login
17 credentials, IP addresses, and other sensitive personal information including analytics and cookies.⁸⁵

18 89. Defendants aggregate all of this data with the entirety of every internet user’s digital
19 footprint, scraped before ChatGPT was available for use, arming them with the largest corporate
20 collection of personal online information ever amassed. Given Defendants’ ongoing theft, this
21 goldmine of valuable data is growing day by day, and with it, the concomitant risk to millions of
22 consumers.

23 90. Indeed, even more stunning than Defendants’ conversion of the internet for
24

25 ⁸³ Baba Tamim, *OpenAI Changes AI Strategy, Won’t Train ChatGPT on Customer Data, Says*
26 *Sam Altman*, INTERESTING ENG’G (May 6, 2023),
<https://interestingengineering.com/culture/openai-wont-train-chatgpt-on-customer-data>.

27 ⁸⁴ *Terms of Use*, OPENAI, <https://openai.com/policies/terms-of-use> (last updated Mar. 14, 2023).

28 ⁸⁵ *Privacy Policy*, OPENAI <https://openai.com/policies/privacy-policy> (last updated June 23, 2023);
Sarah Moore, *What Does ChatGPT Mean for Healthcare?*, NEWS MED. (Mar. 28, 2023),
<https://www.news-medical.net/health/What-does-ChatGPT-mean-for-Healthcare.aspx>.

1 commercial gain, is they are “entrusting” all this personal information to large language models
 2 and unpredictable human-like “bots”, while openly acknowledging that even they “don’t understand
 3 how it works.”⁸⁶ In the words of Mr. Altman himself, “the scary part” is that OpenAI’s act of
 4 “putting this lever into the world *will for sure have unpredictable consequences.*”⁸⁷ Dr. Yoshua
 5 Benigo, one of the three scientists who spent decades developing the technology that drives systems
 6 like ChatGPT-4, further explained: “Our ability to understand what could go wrong with very
 7 powerful A.I. systems is very weak. . . So we need to be careful.”⁸⁸

8 91. To risk the personal data of millions by incorporating all of it into unpredictable
 9 Products, built on technology that even Defendants and leading scientists do not completely
 10 understand and thus, necessarily cannot safeguard, and *then* to deploy those Products worldwide for
 11 unfettered use, is the very definition of gross negligence.

12 **E. Microsoft Pushes OpenAI’s Economic Dependence Model**

13 92. Although Defendants’ most recent iteration of ChatGPT (GPT-4) was only recently
 14 released, Defendants have successfully encouraged and injected OpenAI’s products into virtually
 15 every sector—from academia to healthcare. Instead of ensuring its safe launch of the AI models,
 16 Defendants recklessly began deploying the Products into every sector following the economic
 17 dependence model.

18 93. Microsoft has led the charge on the rapid proliferation of ChatGPT throughout the
 19 modern suite of technological applications—integrating the ChatGPT language model into almost
 20 all of its cardinal products and services,⁸⁹ thereby elevating the dangers of data misuse to
 21 unprecedented heights. Microsoft CEO Satya Nadella has indicated that the company plans to
 22

23 ⁸⁶ Jan Leike (@janleike), TWITTER (May 17, 2023, 10:56 AM),
<https://twitter.com/janleike/status/1636788627735736321>.

24 ⁸⁷ Edward Felsenthal & Billy Perrigo, *OpenAI CEO Sam Altman Is Pushing Past Doubts on*
Artificial Intelligence, TIME MAG. (June 21, 2023), [https://time.com/collection/time100-](https://time.com/collection/time100-companies-2023/6284870/openai-disrupters/)
 25 [companies-2023/6284870/openai-disrupters/](https://time.com/collection/time100-companies-2023/6284870/openai-disrupters/) (emphasis added).

26 ⁸⁸ Cade Metz, *What Exactly Are the Dangers Posed By A.I.?*, THE N.Y. TIMES (May 7, 2023),
<https://www.nytimes.com/2023/05/01/technology/ai-problems-danger-chatgpt.html>.

27 ⁸⁹ These services include Bing, GitHub, Teams, and Viva Sales, among others. *See* Bernard Marr,
Microsoft’s Plan to Infuse AI and ChatGPT Into Everything, FORBES (Mar. 6, 2022),
 28 [https://www.forbes.com/sites/bernardmarr/2023/03/06/microsofts-plan-to-infuse-ai-and-chatgpt-](https://www.forbes.com/sites/bernardmarr/2023/03/06/microsofts-plan-to-infuse-ai-and-chatgpt-into-everything/?sh=1adfd46653fc)
[into-everything/?sh=1adfd46653fc](https://www.forbes.com/sites/bernardmarr/2023/03/06/microsofts-plan-to-infuse-ai-and-chatgpt-into-everything/?sh=1adfd46653fc).

1 introduce AI into the remainder of its products in the future.⁹⁰

2 94. ChatGPT is integrated into Microsoft’s search engine, Bing, which has approximately
3 100 million daily active users. ChatGPT has also been integrated into the interface of Microsoft’s
4 flagship communication and collaboration platform, Microsoft Teams, which has 250 million
5 monthly active users.

6 95. Microsoft has also integrated the language model within its digital assistant platform,
7 Cortana, which has an average of 141 million monthly active users.

8 96. Finally, within the Microsoft Dynamics 365 ecosystem, ChatGPT has been employed
9 to power AI-driven customer service chatbots. This has enabled the chatbots to understand and
10 respond to customer queries in a highly human-like manner, thereby significantly increasing the
11 extent of information collected and thus, reducing the need for human intervention in support cases.

12 97. In a real sense, OpenAI now acts as a data scavenging company for Microsoft and
13 provides Microsoft with ChatGPT User and Non-User data belonging to millions of individuals.⁹¹

14 98. The integration of ChatGPT technology into Microsoft’s primary products
15 significantly magnifies existing data privacy concerns. This move effectively enables the collection
16 of consumer information across a wide array of systems and platforms, encompassing a
17 comprehensive range of user interactions. The resultant collation of expansive consumer data
18 contributes to the construction of extensive user profiles.

19 99. This scope of data collection, coupled with user profiling, poses significant potential
20 risks. These risks extend not just to potential breaches of data privacy regulations, but also to the
21 erosion of consumer trust and the potential for misuse of sensitive information.

22 100. Rather than acknowledging these risks and taking steps to mitigate them, Microsoft
23 has laid off its entire “Responsible AI team,” the 10,000 employees within Microsoft’s ethics and
24 society group who were responsible for ensuring that ethical AI principles drive product design.⁹²
25 As one technology news outlet notes, “Data privacy, storage, or usage are probably just fluff talk

26 ⁹⁰ *Id.* (“Every product of Microsoft will have some of the same AI capabilities to completely
27 transform the product.”).

28 ⁹¹ Pandey, *supra* note 30.

⁹² Poulomi Chatterjee, *Why Responsible AI is Just Fluff Talk for Microsoft, Others*, AIM (Mar. 18, 2023), <https://analyticsindiamag.com/why-responsible-ai-is-just-fluff-talk-for-microsoft-others/>.

1 for . . . [Microsoft] anyway.”⁹³

2 101. Other companies have rushed to keep pace, emulating Microsoft by pushing the
3 Products into nearly every conceivable application and service in the past six months of
4 development. As a result, GPT-4 has been integrated into hundreds of applications and platforms
5 over various industries.⁹⁴ According to a Gartner study, the commercial use of AI has increased
6 270% in the last 4 years, with 37% of businesses now using some form of AI technology. By other
7 accounts, the scale of commercial AI is even greater.

8 102. More specifically, AI in general, and OpenAI in particular, is now partnering with an
9 extraordinary number of influential organizations, spreading across the internet completely
10 unchecked.⁹⁵ This has seemingly happened overnight. It was just over six months ago that ChatGPT
11 was released to the public.⁹⁶ In that short span of time, OpenAI integrated with the following major
12 corporations, to name just a few: Snapchat,⁹⁷ Amazon, Microsoft, Expedia, Instacart, Google,
13 BuzzFeed, KAYAK, Shutterstock, Zillow, Wolfram, as well as countless others⁹⁸—including

14 _____
15 ⁹³ Pandey, *supra* note 30

16 ⁹⁴ Bergur Thormundsson, *Amount of Companies Using ChatGPT in their Business Function in*
17 *2023, By Industry*, STATISTA (May 15, 2023),
18 <https://www.statista.com/statistics/1384323/industries-using-chatgpt-in-business/>.

19 ⁹⁵ Beth Floyd, *ChatGPT Plugins*, ROE DIGIT. (May 5, 2023), [https://roedigital.com/ChatGPT-](https://roedigital.com/ChatGPT-plugins/)
20 [plugins/](https://roedigital.com/ChatGPT-plugins/).

21 ⁹⁶ Alyssa Stringer & Kyle Wiggers, *ChatGPT: Everything You Need to Know About the AI-*
22 *Powered Chatbot*, TECHCRUNCH (May 3, 2023), [https://techcrunch.com/2023/05/03/chatgpt-](https://techcrunch.com/2023/05/03/chatgpt-everything-you-need-to-know-about-the-ai-powered-chatbot/?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_sig=AQAAAA-Ab2tIJ3WAdxAd5xb2pWmCPSFqzTyqRmMHEOaaOXsH04KD_DgCLfExvNPrnVX4ioR-uMFVQjAawiyhp5m21A3SqmsPYHv2yHSgfiIdjokmMe981-hq51XH5pWxCfLZOOWwf2wlvK3MnVewrZk4MRmPRAC8ArJXbegg6dnL2-f)
23 [everything-you-need-to-know-about-the-ai-powered-](https://techcrunch.com/2023/05/03/chatgpt-everything-you-need-to-know-about-the-ai-powered-chatbot/?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_sig=AQAAAA-Ab2tIJ3WAdxAd5xb2pWmCPSFqzTyqRmMHEOaaOXsH04KD_DgCLfExvNPrnVX4ioR-uMFVQjAawiyhp5m21A3SqmsPYHv2yHSgfiIdjokmMe981-hq51XH5pWxCfLZOOWwf2wlvK3MnVewrZk4MRmPRAC8ArJXbegg6dnL2-f)
24 [chatbot/?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_](https://techcrunch.com/2023/05/03/chatgpt-everything-you-need-to-know-about-the-ai-powered-chatbot/?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_sig=AQAAAA-Ab2tIJ3WAdxAd5xb2pWmCPSFqzTyqRmMHEOaaOXsH04KD_DgCLfExvNPrnVX4ioR-uMFVQjAawiyhp5m21A3SqmsPYHv2yHSgfiIdjokmMe981-hq51XH5pWxCfLZOOWwf2wlvK3MnVewrZk4MRmPRAC8ArJXbegg6dnL2-f)
25 [_sig=AQAAAA-](https://techcrunch.com/2023/05/03/chatgpt-everything-you-need-to-know-about-the-ai-powered-chatbot/?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_sig=AQAAAA-Ab2tIJ3WAdxAd5xb2pWmCPSFqzTyqRmMHEOaaOXsH04KD_DgCLfExvNPrnVX4ioR-uMFVQjAawiyhp5m21A3SqmsPYHv2yHSgfiIdjokmMe981-hq51XH5pWxCfLZOOWwf2wlvK3MnVewrZk4MRmPRAC8ArJXbegg6dnL2-f)
26 [Ab2tIJ3WAdxAd5xb2pWmCPSFqzTyqRmMHEOaaOXsH04KD_DgCLfExvNPrnVX4ioR-](https://techcrunch.com/2023/05/03/chatgpt-everything-you-need-to-know-about-the-ai-powered-chatbot/?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_sig=AQAAAA-Ab2tIJ3WAdxAd5xb2pWmCPSFqzTyqRmMHEOaaOXsH04KD_DgCLfExvNPrnVX4ioR-uMFVQjAawiyhp5m21A3SqmsPYHv2yHSgfiIdjokmMe981-hq51XH5pWxCfLZOOWwf2wlvK3MnVewrZk4MRmPRAC8ArJXbegg6dnL2-f)
27 [uMFVQjAawiyhp5m21A3SqmsPYHv2yHSgfiIdjokmMe981-](https://techcrunch.com/2023/05/03/chatgpt-everything-you-need-to-know-about-the-ai-powered-chatbot/?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_sig=AQAAAA-Ab2tIJ3WAdxAd5xb2pWmCPSFqzTyqRmMHEOaaOXsH04KD_DgCLfExvNPrnVX4ioR-uMFVQjAawiyhp5m21A3SqmsPYHv2yHSgfiIdjokmMe981-hq51XH5pWxCfLZOOWwf2wlvK3MnVewrZk4MRmPRAC8ArJXbegg6dnL2-f)
28 [hq51XH5pWxCfLZOOWwf2wlvK3MnVewrZk4MRmPRAC8ArJXbegg6dnL2-f](https://techcrunch.com/2023/05/03/chatgpt-everything-you-need-to-know-about-the-ai-powered-chatbot/?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_sig=AQAAAA-Ab2tIJ3WAdxAd5xb2pWmCPSFqzTyqRmMHEOaaOXsH04KD_DgCLfExvNPrnVX4ioR-uMFVQjAawiyhp5m21A3SqmsPYHv2yHSgfiIdjokmMe981-hq51XH5pWxCfLZOOWwf2wlvK3MnVewrZk4MRmPRAC8ArJXbegg6dnL2-f).

⁹⁷ Snapchat recently released “My AI,” a ChatGPT-fueled chatbot feature open to all Snapchat
users. See Alex Hern, *Snapchat Making AI Chatbot Similar to ChatGPT Available to Every User*,
THE GUARDIAN (Apr. 19, 2023), [https://www.theguardian.com/technology/2023/apr/19/snapchat-](https://www.theguardian.com/technology/2023/apr/19/snapchat-making-ai-chatbot-similar-to-chatgpt-available-to-every-user)
making-ai-chatbot-similar-to-chatgpt-available-to-every-user. My AI now appears for Snapchat
users as a contact in their social network, allowing users to ask it questions, have back and forth
conversations, ask it to generate creative content, and much more. *Id.*

⁹⁸ Floyd, *supra* note 95; Silvia Pellegrino, *Which Companies Have Partnered With OpenAI*,
TECHMONITOR (Jan. 18, 2023), [https://techmonitor.ai/technology/which-companies-have-](https://techmonitor.ai/technology/which-companies-have-partnered-with-openai)
partnered-with-openai; Asif Iqbal, *OpenAI’s Collaborations: Pushing the Boundaries of AI in*
Various Sectors, LINKEDIN (Mar. 12, 2023), [https://www.linkedin.com/pulse/openai-](https://www.linkedin.com/pulse/openai-collaborations-pushing-boundaries-ai-various-sectors-iqbal/)
collaborations-pushing-boundaries-ai-various-sectors-iqbal/.

1 everything from pioneering drug treatments in the health sector (Pfizer)⁹⁹ to optimizing dating
2 applications (OkCupid).¹⁰⁰ At this point, it might be easier to list the companies that have not
3 partnered with OpenAI, or that are not investing in their own AI solutions.

4 103. As is clear, OpenAI has exploded outwards in every direction within the past few
5 months and is swiftly morphing into something intimately connected with people in nearly every
6 aspect of their day-to-day lives. There is no check or boundary on this expansion, which seems to
7 progress rapidly every single day.

8 **II. Risks from Unchecked AI Proliferation**

9 **A. The International Community Agrees that Unchecked & Lawless AI** 10 **Proliferation Poses an Existential Threat**

11 104. The unregulated development of AI technology has led to the creation of powerful
12 tools being used to manipulate public opinion, spread false information, and undermine democratic
13 institutions. Further development of such powerful tools will supercharge the dissemination of
14 propaganda, the amplification of extremist voices, and the influencing of elections based on
15 undetectable falsehoods.

16 105. The United States has been particularly affected by the rapid development of AI
17 technology, as the absence of effective regulations has accelerated the proliferation of
18 unaccountable and untrustworthy AI tools. Even the White House has acknowledged that AI
19 presents “the most complicated tech policy discussion possibly that [the country] has ever had.”¹⁰¹

20 ***“I am confident AI will be used by bad actors, and yes it will cause real damage.”***¹⁰² -

21 *Michael Schwarz, Microsoft’s Chief Economist*

22 ***“If law and due process are absent from this field, we are essentially paving the way to a***

23 ⁹⁹ Iqbal, *supra* note 98 (“In 2020, OpenAI announced a collaboration with drug manufacturer,
24 Pfizer, to develop new AI technologies for drug discovery.”).

25 ¹⁰⁰ Danni Button, *ChatGPT Poses Danger for Online Dating Apps*, THE STREET (Feb. 15, 2023),
<https://www.thestreet.com/social-media/chatgpt-poses-dangers-for-online-dating-apps>.

26 ¹⁰¹ Ben Wershkul & Alexandra Garfinkle, *White House bringing Google, Microsoft CEOs*
together for ‘frank discussion’ of AI, YAHOO! FIN. (May 4, 2023),

27 <https://www.aol.com/finance/white-house-bringing-alphabet-microsoft-164428066.html>.

28 ¹⁰² Bryce Baschuk, *Microsoft Economist Warns Bad Actors Will Use AI to Cause Damage*, MSN
(May 3, 2023), <https://www.msn.com/en-us/money/other/ai-will-cause-real-damage-microsoft-chief-economist-warns/ar-AA1aFslV>.

1 *new feudal order of unaccountable reputational intermediaries.” - Professors Danielle*
 2 *Keats Citron and Frank Pasquale at 2023 Geneva Conference.*¹⁰³

3 *AI technology is so powerful that it even has the potential to “allow an evil country,*
 4 *competitor to come in and screw up our democracy.”*¹⁰⁴ - Eric Schmidt, Former Google
 5 *CEO and Chairman at the 2023 Milken Global Conference.*

6 106. In a report addressed to the American public in 2021, Eric Schmidt and Robert Work,
 7 the chair and vice chair of the National Security Commission on Artificial Intelligence (“NSCAI”),
 8 noted that “Americans have not yet grappled with just how profoundly the artificial intelligence
 9 revolution will impact our economy, national security, and welfare. Much remains to be learned
 10 about the power and limits of AI technologies. Nevertheless, **big decisions need to be made now** . . . to
 11 defend against the malignant uses of AI.”¹⁰⁵

12 107. The NSCAI report highlights the consequences associated with the unregulated
 13 development of AI, emphasizing the unique risks to human rights, privacy, and personal autonomy.
 14 Further, the report notes the urgency of establishing comprehensive privacy frameworks and
 15 regulations that strike a balance between protecting individuals’ privacy rights and enabling AI
 16 advancements.

17 108. On March 30, 2023, a new complaint was filed to the Federal Trade Commission
 18 (“FTC”), urging the agency to investigate OpenAI and suspend its commercial deployment of large
 19 language models, including its latest iteration of the popular tool ChatGPT.¹⁰⁶ The complaint notes
 20 that the use of AI should be “transparent, explainable, fair, and empirically sound while fostering
 21 accountability.”¹⁰⁷ None of the Products satisfy these requirements.

22 109. The significance of harm facing our society is in fact so imminent that Geoffrey
 23

24 ¹⁰³ *EPIC AI Rulemaking Petition*, EPIC, <https://epic.org/documents/epic-ai-rulemaking-petition/>
 25 (last visited June 27, 2023).

26 ¹⁰⁴ Wershkul, *supra* note 101.

27 ¹⁰⁵ Eric Schmidt & Bob Work, *Letter from the Chair and Vice Chair*, NAT’L. SEC. COMM’N. ON
 28 A.I., (2021), <https://reports.nsc.ai.gov/final-report/chair-and-vice-chair-letter>.

29 ¹⁰⁶ Federal Trade Commission, *In the matter of OpenAI, Inc.*, FED. TRADE. COMM’N. (Mar. 30,
 30 2023), [https://cdn.arstechnica.net/wp-content/uploads/2023/03/CAIDP-FTC-Complaint-OpenAI-
 31 GPT-033023.pdf](https://cdn.arstechnica.net/wp-content/uploads/2023/03/CAIDP-FTC-Complaint-OpenAI-GPT-033023.pdf).

32 ¹⁰⁷ *Id.*

1 Hinton—referenced by many as the “godfather” of AI—quit his job at Google where he had worked
 2 for more than a decade, becoming one of the most respected voices in the field, so he could freely
 3 speak out about the dangers associated with the rapid, uncontrolled development and release of AI
 4 to our society.

5 110. Dr. Hinton’s journey from AI groundbreaker to AI whistleblower marks a remarkable
 6 moment for the AI technology industry at perhaps its most important inflection point in decades.
 7 Industry leaders believe the new A.I. systems could be as important but yet as catastrophic as the
 8 development of nuclear weapons.

9 111. After OpenAI released ChatGPT in March, more than 1,000 technology leaders and
 10 researchers signed an open letter calling for a six-month moratorium on the development of new
 11 systems because A.I. technologies pose “profound risks to society and humanity.”¹⁰⁸

12 112. Several days later, 19 current and former leaders of the Association for the
 13 Advancement of Artificial Intelligence, a 40-year-old academic society, released their own letter
 14 warning of the risks of A.I. That group included Eric Horvitz, chief scientific officer at Microsoft,
 15 which has deployed OpenAI’s technology across a wide range of products, including its Bing search
 16 engine.¹⁰⁹

17 113. The Letter, issued by the Future of Life Institute, states:

18 **Powerful AI systems should be developed only once we are confident**
 19 **that their effects will be positive and their risks will be manageable . . .**
 20 **we call on all AI labs to immediately pause for at least 6 months the**
 21 **training of AI systems more powerful than GPT-4.** AI research and
 22 development should be refocused on making today’s powerful, state-of-the-
 art systems more accurate, safe, interpretable, transparent, robust, aligned,
 trustworthy, and loyal.¹¹⁰

23 114. The Letter continues: “In parallel, AI developers must work with policymakers to
 24 dramatically accelerate development of robust AI governance systems. These should at a minimum
 25 include new and capable regulatory authorities dedicated to AI; . . .”¹¹¹

26 ¹⁰⁸ *The ‘Godfather of A.I.’ Leaves Google and Warns of Danger Ahead*, DNYUZ (May 1, 2023),
<https://dnyuz.com/2023/05/01/the-godfather-of-a-i-leaves-google-and-warns-of-danger-ahead/>.

27 ¹⁰⁹ *Id.*

28 ¹¹⁰ *Pause Giant AI Experiments: An Open Letter*, FUTURE OF LIFE INST. (Mar. 29, 2023),
<https://futureoflife.org/open-letter/pause-giant-ai-experiments/> (emphasis in the original).

¹¹¹ *Id.*

1 115. Generative AI models are unusual consumer products because they exhibit behaviors
2 that may not have been previously identified by the company that released them for sale. OpenAI
3 acknowledged the risk of “Emergent Risky Behavior” and nonetheless chose to go forward with the
4 commercial release of ChatGPT. As OpenAI explained: novel capabilities often emerge in more
5 powerful models. Some that are particularly concerning are the ability to create and act on long-
6 term plans, to accrue power and resources (“power-seeking”), and to exhibit behavior that is
7 increasingly “agentic.”¹¹²

8 116. In February 2020, a petition with the Federal Trade Commission called on the FTC to
9 conduct rulemaking for the use of artificial intelligence in commerce. “Given the scale of
10 commercial AI use, the rapid pace of AI development, and the very real consequences of AI-enabled
11 decision-making for consumers, [courts] should immediately initiate a rulemaking to define and
12 prevent consumer harms resulting from AI.”¹¹³

13 117. Multiple sources have called on the FTC to enforce the AI standards established in
14 the OECD AI Principles, the OMB AI Guidance, and the Universal Guidelines for AI. Several FTC
15 Commissioners have already acknowledged the FTC’s role in regulating the use of AI.

16 118. The absence of effective AI regulations in the United States has accelerated the spread
17 of unaccountable and untrustworthy AI tools. And the unregulated use of those AI tools has already
18 caused serious harm to consumers, who are increasingly subject to opaque and unprovable decision-
19 making in employment, credit, healthcare, housing, and criminal justice.

20 119. Realizing the gravity of potential harm, authorities within European countries took
21 ChatGPT offline in Italy in April after the country’s data protection authority temporarily banned
22 the chatbot and launched a probe into the artificial intelligence application’s suspected breach of
23

24
25 ¹¹² Dennis Layton, *GPT-4 – Some First Impressions*, LINKEDIN (Mar. 15, 2023),
26 <https://www.linkedin.com/pulse/gpt-4-some-first-impressions-dennis-layton> (“Agentic in this
27 context does not intend to humanize language models or refer to sentience but rather refers to
28 systems characterized by the ability to, e.g., accomplish goals which may not have been concretely
specified and which have not appeared in training; focus on achieving specific, quantifiable
objectives; and [engage in] long-term planning.”).

¹¹³ *EPIC AI Rulemaking Petition*, *supra* note 103.

1 privacy rules.¹¹⁴

2 120. Italian authorities stated that ChatGPT has an “absence of any legal basis that justifies
3 the massive collection and storage of personal data” to “train” the chatbot.¹¹⁵ Further, they accused
4 Defendant OpenAI of failing to check the age of ChatGPT’s users to ensure they are aged 13 or
5 above.¹¹⁶

6 121. Subsequently, Defendant OpenAI agreed to offer specific tools to verify Users’ ages
7 in Italy upon sign-up, but yet continues to enable unverified access in the United States to illegally
8 collect the personal data of minors. Defendant OpenAI also said that it would provide greater
9 visibility of its privacy policy and user content opt-out form, creating a new form for European
10 Union users to exercise their right to object to its use of personal data to train its models. The form
11 requires people who want to opt out to provide detailed personal information, including evidence of
12 data processing via relevant prompts. However, despite consumers’ established privacy rights to be
13 “forgotten,” Defendants cannot effectively extract individuals’ information from the Products once
14 the AI is trained on such information.¹¹⁷

15 122. Italy was the first western European country to curb ChatGPT, but its rapid
16 development has attracted attention from lawmakers and regulators in several countries. A
17 committee of European Union lawmakers agreed on new rules that would force companies
18 deploying generative AI tools, such as ChatGPT, to disclose any copyrighted material used to
19 develop their systems.¹¹⁸

21 ¹¹⁴ Supantha Mukherjee & Giselda Vagnoni, *Italy Restores ChatGPT After OpenAI Responds to*
22 *Regulator*, YAHOO! (Apr. 28, 2023), [https://finance.yahoo.com/news/chatgpt-available-again-
users-italy-163139143.html](https://finance.yahoo.com/news/chatgpt-available-again-users-italy-163139143.html).

23 ¹¹⁵ Elvira Pollina & Supantha Mukherjee, *Italy Curbs ChatGPT, Starts Probe Over Privacy*
24 *Concerns*, REUTERS (Mar. 31, 2023), [https://www.reuters.com/technology/italy-data-protection-
agency-opens-chatgpt-probe-privacy-concerns-2023-03-31/](https://www.reuters.com/technology/italy-data-protection-agency-opens-chatgpt-probe-privacy-concerns-2023-03-31/).

25 ¹¹⁶ *Id.*

26 ¹¹⁷ *ChatGPT and Education*, CNT. FOR INNOVATIVE TEACHING AND LEARNING,
27 <https://www.niu.edu/citl/resources/guides/chatgpt-and-education.shtml>, (last visited June 26,
2023) (“the prompts that you input into ChatGPT cannot be deleted. If you, or your students, were
to ask ChatGPT about sensitive or controversial topics, this data cannot be removed.”).

28 ¹¹⁸ Supantha Mukherjee & Giselda Vagnoni, *Italy Restores CHATGPT after OpenAI Responds to*
Regulator, SRN NEWS (Apr. 28, 2023), [srnnews.com/italy-restores-chatgpt-after-openai-responds-
to-regulator-2/](https://srnnews.com/italy-restores-chatgpt-after-openai-responds-to-regulator-2/).

1 123. Data authorities from around the world remain concerned, specifically, with “the lack
2 of legal basis underpinning the massive collection, use and disclosure of personal information in
3 order to train the ChatGPT algorithms on which the platform relies” and the “cornerstone privacy
4 issue” at the heart of this Action: ChatGPT’s “use of web scraping and the collection of personal
5 information without consent.”¹¹⁹

6 124. In short, the message is consistent from informed business, nonprofit, and technology
7 thought leaders; industrialists; scientists; world leaders; regulators; and governments around the
8 globe: The proliferation of AI—including Defendants’ products—pose an existential threat if not
9 constrained by the reasonable guardrails of our laws and societal mores. Defendants’ business and
10 scraping practices raise fundamentally important legal and ethical questions that must also be
11 addressed. Enforcing the law will not amount to stifling AI innovation, but rather a safe and just AI
12 future for all.

13 **B. Overview of Risks**

14 125. The following is a brief, non-exhaustive list of ongoing harms and critical legal threats
15 the Products pose to everyday Americans, including Plaintiffs and the Proposed Class Members.

16 ***1. Massive Privacy Violations***

17 126. In today’s vast, interconnected digital landscape, privacy can appear to be more of an
18 illusion, but it is still a guaranteed right. In violation of this right, the Products operate as an all-
19 seeing online platform, tracking our every move: each click, each site visit, each chat—not allowing
20 anything to escape its relentless scrutiny. Internet users’ interactions, seemingly innocuous, are

21
22 ¹¹⁹ Roland Hung, *AI Technology and Privacy: Canadian Privacy Commissioner Launches*
23 *Investigation into ChatGPT*, TORKIN MANES (Apr. 24, 2023), [https://www.torkinmanes.com/our-
resources/publications-presentations/publication/ai-technology-and-privacy-canadian-privacy-
commissioner-launches-investigation-into-chatgpt](https://www.torkinmanes.com/our-resources/publications-presentations/publication/ai-technology-and-privacy-canadian-privacy-commissioner-launches-investigation-into-chatgpt) (detailing the “privacy concerns with the use of
24 ChatGPT” that have been raised worldwide). *See also* Heinrich Long, *Authorities Press OpenAI to*
25 *Disclose How ChatGPT Input Is Used*, RESTORE PRIV. (June 9, 2023),
26 <https://restoreprivacy.com/authorities-press-openai-to-disclose-how-chatgpt-input-is-used/>
27 (discussing worldwide investigations, including the latest inquiry from Dutch data protection
28 authorities who “want[] to know, among other things, how OpenAI handles personal data when
training the underlying system. The[y...] want[] to know from OpenAI whether people’s
questions are used to train the algorithm, and if so, in what way. The[y...] also ha[ve] questions
about the way in which OpenAI collects and uses personal data from the internet.”).

1 aggregated, filtered, and compiled by Defendants, rendering the concept of privacy virtually non-
2 existent. Even information deemed private or intended for a restricted audience does not escape
3 surveillance.

4 127. The massive, unparalleled collection and tracking of users’ personal information by
5 Defendants endangers individuals’ privacy and security to an incalculable degree. This information
6 can be exploited and used to perpetrate identity theft, financial fraud, extortion, and other malicious
7 purposes. It can also be employed to target vulnerable individuals with predatory advertising,
8 algorithmic discrimination, and other unethical and harmful acts.

9 128. The collection and use of this data raises concerns about user privacy and the potential
10 misuse of personal information. For example, every iota of Users’ activity is tracked and monitored.
11 By analyzing this data using algorithms and machine learning techniques, Defendants can develop
12 a chillingly detailed understanding of users’ behavior patterns, preferences, and interests—creating
13 an entirely new meaning to the term “invasive.”

14 129. Several studies confirm that the collection and disclosure of sensitive information
15 from millions of individuals, as Defendants have done here, violates established expectations of
16 privacy based on long-standing social norms. Privacy polls and studies uniformly show that the
17 overwhelming majority of Americans consider one of the most important privacy rights to be the
18 need for an individual’s affirmative consent before a company collects and shares its customers’
19 data.

20 130. For example, a recent study by Consumer Reports reveals that 92% of Americans
21 believe that internet companies and websites should be required to obtain consent before selling or
22 sharing consumers’ data, and that internet companies and websites should be required to provide
23 consumers with a complete list of the data that has been collected about them.¹²⁰ Moreover,
24 according to a study by Pew Research Center, a majority of Americans, approximately 79%, are
25

26
27
28

¹²⁰ *Consumers Less Confident About Healthcare, Data Privacy, and Car Safety, New Survey Finds*, CONSUMER REPS. (May 11, 2017), <https://www.consumerreports.org/consumer-reports/consumers-less-confident-about-healthcare-data-privacy-and-car-safety/>.

1 concerned about how companies collect data about them.¹²¹

2 131. Users act consistently with these privacy preferences. Following a new rollout of the
3 iPhone operating software—which asks users for clear, affirmative consent before allowing
4 companies to track users—85% of worldwide users and 94% of U.S. users chose not to share data
5 when prompted.¹²² The Products’ Users do not have that option, and do not understand the full extent
6 of Defendants’ data collection and use of their personal data.

7 132. While the reams of personal information that Defendants collect on Users can be used
8 to provide personalized and targeted responses, it can also be used for exceedingly nefarious
9 purposes, such as tracking, surveillance, and crime. For example, if ChatGPT has access to a User’s
10 browsing history, search queries, and geolocation, and combines this information with what
11 Defendant OpenAI has secretly scraped from the internet, Defendants could build a detailed profile
12 of Users’ behavior patterns, including but not limited to where they go, what they do, with whom
13 they interact, and what their interests and habits are. This level of surveillance and monitoring raises
14 vital ethical and legal questions about privacy, consent, and the use of personal data. It is crucial for
15 users to be aware of how their data is being collected and used, and to have control over how their
16 information is shared and used by advertisers and other entities.

17 133. The concern about collecting and sharing information is compounded by the reality
18 that this information may include particularly sensitive information such as medical records or
19 information about minors. Increasingly, companies like Defendants “are harnessing and collecting
20 multiple typologies of children’s data and have the potential to store a plurality of data traces under
21 unique ID profiles.”¹²³

22 134. Given ChatGPT’s ability to generate human-like understanding and responses, there
23

24 ¹²¹ Brooke Auxier et al., *Americans and Privacy: Concerned, Confused, and Feeling Lack of*
25 *Control over Their Personal Information*, PEW RSCH. CTR. (Nov. 15, 2019),
26 [https://www.pewresearch.org/internet/2019/11/15/americans-and-privacy-concerned-confused-](https://www.pewresearch.org/internet/2019/11/15/americans-and-privacy-concerned-confused-and-feeling-lack-of-control-over-their-personal-information/)
[and-feeling-lack-of-control-over-their-personal-information/](https://www.pewresearch.org/internet/2019/11/15/americans-and-privacy-concerned-confused-and-feeling-lack-of-control-over-their-personal-information/).

27 ¹²² Margaret Taylor, *How Apple Screwed Facebook*, WIRED (May 19, 2021, 6:00 AM),
<https://www.wired.co.uk/article/apple-ios14-facebook>.

28 ¹²³ Veronica Barassi, *Tech Companies Are Profiling Us from Before Birth*, THE MIT PRESS
READER, (Jan. 14, 2021), [https://thereader.mitpress.mit.edu/tech-companies-are-profiling-us-from-](https://thereader.mitpress.mit.edu/tech-companies-are-profiling-us-from-before-birth/)
[before-birth/](https://thereader.mitpress.mit.edu/tech-companies-are-profiling-us-from-before-birth/).

1 is a high likelihood that users might share (and already are sharing) their private health information
2 while interacting with the model, by asking health-related questions or discussing their medical
3 history, symptoms, or conditions. Moreover, this information can be logged and reviewed as part of
4 ongoing efforts to “train,” improve and monitor each model’s performance.

5 135. However, beyond these seemingly innocuous interactions with the AI, healthcare
6 industry providers are beginning to integrate ChatGPT in order to “revolutionize healthcare” while
7 undermining the confidentiality of individuals’ personal data, which would be transmitted using
8 ChatGPT and continuing to train Defendants’ AI at the patients’ expense.¹²⁴ While this technology
9 could provide benefits, the risks associated with its implementation are drastic, from cybercrime,
10 misinformation and misdiagnosis, lack of empathy and experience, and bias¹²⁵ to the existential risk,
11 of which Altman has repeatedly warned.

12 136. ***Established Privacy Rights to be “Forgotten” Violated.*** Compounding this massive
13 invasion of privacy, OpenAI offers no *effective* procedures at this time for individuals to request for
14 their information/training data to be deleted. Instead, OpenAI simply provides an email address that
15 consumers can contact if they would like to have their information removed. But this “option” is
16 illusory. Regardless of whether individuals can technically request for ChatGPT to remove their
17 data, it is not possible to do so completely, because Defendants train ChatGPT on individual inputs,
18 personal information, and other user and nonuser data, which Defendants cannot reliably and fully
19 extract from its trained AI systems any more than a person can “unlearn” the math they learned in
20 sixth grade.

21 137. An AI researcher with privacy and cybersecurity firm AVG explains, “People are
22 furious that data is being used without their permission. . . Sometimes, some people have deleted
23 the[ir] [online] data but since the language model has already used them, the data is there forever.
24
25

26 ¹²⁴ Naomi Diaz, *6 Hospitals, Health Systems Checking Out ChatGPT*, BECKERS HEALTHCARE
27 (June 2, 2023), <https://www.beckershospitalreview.com/innovation/4-hospitals-health-systems-testing-out-chatgpt.html>.

28 ¹²⁵ Ethan Popowitz, *ChatGPT: Friend or Foe?*, DEFINITIVE HEALTHCARE,
<https://www.definitivehc.com/blog/chatgpt> (last visited June 27, 2023).

1 They don't know how to delete the data.”¹²⁶

2 138. Likewise, some companies have banned or limited ChatGPT use because they are
3 “worried that anything uploaded to AI platforms like OpenAI’s ChatGPT or Google’s Bard will
4 [also] get *stored* on those companies’ servers, *with no way to access or delete the information.*”¹²⁷

5 139. The “right to be forgotten”—i.e., the right to request that a business delete the personal
6 information that it holds about you—is guaranteed to California residents under the California
7 Consumer Privacy Act of 2018 (“CCPA”). Given how the technology works, OpenAI is not
8 compliant with these requirements.¹²⁸

9 **2. AI-Fueled Misinformation Campaigns, Targeted Attacks, Sex Crimes, and**
10 **Bias**

11 140. **Misinformation, Deepfakes, Clones, Scams, and Blackmail:** The use of the Products
12 facilitates the spreading of false or misleading information, even without “misuse.” That is because
13 a *feature* (known defect) of ChatGPT’s *regular use* is the inventing of false information, including
14 potentially defamatory information about individuals. Even the “improved” version (GPT4) “makes
15 stuff up” and “may generated text that is completely false.”¹²⁹

16 141. One high-profile example involves a US law professor, Jonathan Turley, who
17 ChatGPT falsely accused of sexually harassing one of his students, even providing a “source” for
18 the purported crime via a news article that it invented.¹³⁰ Defendants call this “hallucination,” but
19

20 ¹²⁶ *Is ChatGPT’s use of people’s data even legal?*, AVG, <https://www.avg.com/en/signal/chatgpt-data-use-legal?> (last visited June 27, 2023).

21 ¹²⁷ Felicity Nelson, *Many Companies are Banning ChatGPT. This is Why*, SCI. ALERT (June 16,
22 2023), <https://www.sciencealert.com/many-companies-are-banning-chatgpt-this-is-why> (emphasis
23 added). Microsoft has itself directed employees not to share sensitive data with ChatGPT “in case
24 it’s used for future AI training models” Diamond Naga Siu, *Microsoft is chill with employees
using ChatGPT — just don’t share ‘sensitive data’ with it*, YAHOO! NEWS (Feb. 1, 2023),
<https://news.yahoo.com/microsoft-chill-employees-using-chatgpt-114000174.html?guccounter=1>.

25 ¹²⁸ See, e.g., Alexa Johnson-Gomez, *A “Living” AI: How ChatGPT Raises Novel Data Privacy
Issues*, MINN. J. OF L., SCI. & TECH. BLOG (Feb. 6, 2023), <https://mjlst.lib.umn.edu/2023/02/06/a-living-ai-how-chatgpt-raises-novel-data-privacy-issues/> (dismissing purported compliance with
26 CCPA as “in name only” given how the data is used as part of machine learning model).

27 ¹²⁹ Cade Metz, *10 Ways GPT-4 is Impressive but Still Flawed*, THE N.Y. TIMES (Mar. 14, 2023),
28 <https://www.nytimes.com/2023/03/14/technology/openai-new-gpt4.html>.

¹³⁰ Hern and Milmo, *supra* note 59.

1 the world knows it as defamation. While Defendants are allegedly “working on” a fix for this
 2 behavior, they continue to push the defective Product worldwide. Naturally, one would expect an
 3 ethical company “for the *benefit* of humanity” not to release such a Product, at all, *unless and until*
 4 it was safeguarded from committing crimes *against* humanity.

5 142. The Cambridge Analytica scandal—in which personal data was allegedly misused to
 6 target individuals with political propaganda and misinformation—is also an instructive cautionary
 7 tale.¹³¹ Cambridge Analytica collected personal data using third-party apps that collected data from
 8 users and their friends. It then used this data to build detailed profiles of individuals, so they could
 9 be targeted with personalized political ads and propaganda. Cambridge Analytica used algorithms
 10 and machine learning techniques to analyze this data, identify patterns in users’ behavior and
 11 preferences, and target those users with specific messages and ads.

12 143. This history highlights the potential dangers of using personal data to build detailed
 13 profiles of individuals, particularly when that data is collected without their knowledge or consent.
 14 It also raises important questions about the ethics of using personal data for political purposes and
 15 the need for greater regulation and oversight of data collection and use.

16 144. Moreover, by allowing the collection, storage, and analysis of a massive amount of
 17 highly individualized, personal data—from audio and photographic data to detailed interests, habits,
 18 and preferences—OpenAI’s technology facilitates the proliferation of video or audio “deepfakes”
 19 and makes them harder to detect.¹³² Simply put, the Products make it easier to create lifelike
 20 audiovisual digital duplicates--digital clones—of real people, which can then be used to spread
 21 misinformation, exploit victims, or even access privileged data.¹³³

22 ¹³¹ Sam Meredith, *Here’s Everything You Need to Know About the Cambridge Analytica Scandal*,
 23 CNBC (Mar. 21, 2018), <https://www.cnbc.com/2018/03/21/facebook-cambridge-analytica-scandal-everything-you-need-to-know.html>. (The Cambridge Analytica scandal involved the
 24 misuse of personal data collected from Facebook users, which was then used to target individuals
 25 with political advertising and propaganda. The scandal highlighted the potential dangers of using
 26 personal data for targeted advertising and the need for greater transparency and accountability in
 the collection and use of personal information.).

27 ¹³² Bibhu Dash & Pawankumar Sharma, *Are ChatGPT and Deepfake Algorithms Endangering the
 Cybersecurity Industry? A Review*, 10(1) I. J. OF ENG’G & APPLIED SCI. (Jan. 2023),
https://www.ijeas.org/download_data/IJEAS1001001.pdf.

28 ¹³³ *Science & Tech Spotlight: Deepfakes*, U.S. GOV’T ACCOUNTABILITY OFF. (Feb. 20, 2020),
<https://www.gao.gov/products/gao-20-379sp>; *see also* Dash & Sharma, *supra* note 132.

1 145. Deepfakes could influence elections, erode public trust, and negatively affect public
2 discourse.¹³⁴ The U.S. Congressional Research Service has further analyzed the risks of deepfakes,
3 explaining that they could be used to “blackmail elected officials or individuals with access to
4 classified information” and “generate inflammatory content [...] intended to radicalize populations,
5 recruit terrorists, or incite violence.”¹³⁵

6 146. In addition to spreading misinformation, criminals have used, and will continue to use
7 this technology to harass, blackmail, extort, coerce, and defraud. Armed with artificial intelligence
8 tools like the ones developed by Defendants, malicious actors can weaponize even the most
9 innocuous publicly available personal information, such as names and photographs, against private
10 individuals.

11 147. For example, the FBI has issued an alert about a particularly despicable form of
12 blackmail currently on the rise that has been largely facilitated by AI like the Products. This scheme,
13 a form of “sextortion,” is perpetrated using artificial intelligence tools and publicly available
14 photographs and videos of private individuals, usually obtained through social media, to create
15 deepfakes containing pornographic content.¹³⁶ The photos or videos are then publicly circulated on
16 social media, public forums, and pornographic websites for the purpose of harassing the victim,
17 causing extreme emotional and psychological distress.¹³⁷

18 148. A malicious actor may also attempt to extract ransom payments, sometimes seeking
19 genuine versions of the subject engaging in the acts depicted in the made up sexually-explicit images
20 and videos, by threatening to share the falsified images or videos with family members, social
21 contacts, or by indiscriminately circulating the content on social media.¹³⁸ The most concerning and
22 egregious aspect of this type of “sextortion” scheme is that the victims include not only non-
23

24
25 ¹³⁴ Kelley M. Saylor & Laurie A. Harris, *Deep Fakes and National Security*, CONG. RSCH. SERV.,
(April 17, 2023), <https://crsreports.congress.gov/product/pdf/if/if11333>.

26 ¹³⁵ *Id.*

27 ¹³⁶ *Public Service Announcement: Malicious Actors Manipulating Photos and Videos to Create
Explicit Content and Sextortion Schemes*, FED. BUREAU OF INVESTIGATION (June 5, 2023),
<https://www.ic3.gov/Media/Y2023/PSA230605>.

28 ¹³⁷ *Id.*

¹³⁸ *Id.*

1 consenting adults, but also minor children.¹³⁹

2 149. **Child Pornography.** Defendants’ Product Dall-E has become a favorite tool for
3 pedophiles, because it requires less technical competence than previous programs used by
4 pedophiles and increases the scale at which images of virtual child pornography can be created.¹⁴⁰
5 In just mere seconds, Dall-E can create realistic images of children performing sex acts.¹⁴¹
6 Thousands of such images have already been detected in dark web forums.¹⁴² In a dark web forum
7 with 3,000 subscribers, 80% of respondents to an internal poll stated “they had used or intended to
8 use AI tools to create child sexual abuse images.”¹⁴³ In such forums, users exchange strategies for
9 thwarting the woefully insufficient purported “safety guardrails” of Dall-E and other AI products,
10 “including by using non-English languages they believe are less vulnerable to suppression or
11 detection.”¹⁴⁴

12 150. Dall-E is a diffusion model, and anyone can access it, generating a realistic image
13 solely by typing a short description of the desired product.¹⁴⁵ This model was trained off billions of
14 images taken, without notice or consent, from the internet, “many of which showed real children
15 and came from photo sites and personal blogs.”¹⁴⁶ Images of actual children are thus the source
16 material for the AI-generated child pornography. In some instances, actual images of existing child
17 pornography were used to train the model and generate further explicit material of already
18 victimized children, thereby victimizing them all over again.¹⁴⁷

19 151. AI-generated child pornography has introduced a slew of other horrendous problems
20 as well. “The flood of images could confound the central tracking system built to block such material
21 from the web because it is designed only to catch known images of abuse, not detect newly generated

22 ¹³⁹ *Id.*

23 ¹⁴⁰ Drew Harwell, *AI-generated Child Sex Images Spawn New Nightmare for the Web*, WASH.
24 POST (June 19, 2023, 7:00 AM),
[https://www.washingtonpost.com/technology/2023/06/19/artificial-intelligence-child-sex-abuse-
25 images/](https://www.washingtonpost.com/technology/2023/06/19/artificial-intelligence-child-sex-abuse-images/).

25 ¹⁴¹ *Id.*

26 ¹⁴² *Id.*

26 ¹⁴³ *Id.*

27 ¹⁴⁴ *Id.*

27 ¹⁴⁵ *Id.*

28 ¹⁴⁶ *Id.*

¹⁴⁷ *Id.*

ones.”¹⁴⁸ Moreover, the monumental task of locating children harmed by the production of child pornography has been bogged down now that agents must now spend time puzzling over whether content is real or virtual.¹⁴⁹ Furthermore, this virtual material is not merely used by pedophiles to supplant real material.¹⁵⁰ AI is also being used to “build [] fake school-age persona[s]” via fabricated selfies, which are incorporated into plots to lure and groom child targets.¹⁵¹

152. Absent the injunctive relief sought in this action, Defendants will continue to not only steal data from unwitting victims, including minors, but arm pedophiles in rapidly generating child pornography at scale and in creating materials that can be strategically used to groom and victimize real children.

153. ***Hate and Bias.*** Continued commercial deployment of the Products also will amplify and entrench the human biases and prejudices reflected in the Products’ sources, which Defendants used without regard to such factors by incorporating and training the Products with content from various extremist websites and by failing to use adequate filtering safeguards.¹⁵²

3. ***Hypercharged Malware Creation***

154. ***Malicious, Mutating, and Virtually Undetectable Code Scripts:*** Malware, or malicious software, are computer programs designed to damage or infiltrate computer systems. Unscrupulous actors deploy malware by embedding them within vulnerabilities in existing internet applications.¹⁵³ The Products guarantee that “malware” prevalence and potency will exponentially increase, posing unprecedented cybersecurity risks on a global scale. That is because the Products

¹⁴⁸ *Id.*

¹⁴⁹ *Id.*

¹⁵⁰ *Id.*

¹⁵¹ *Id.*

¹⁵² Sam Biddle, *The Internet’s New Favorite AI Proposes Torturing Iranians and Surveilling Mosques*, THE INTERCEPT (Dec. 8, 2022), <https://theintercept.com/2022/12/08/openai-chatgpt-ai-bias-ethics/>.

¹⁵³ Fei Xiao et al., *A Novel Malware Classification Method Based on Crucial Behavior*, 2020 MATHEMATICAL PROBS. IN ENG’G. (Mar. 21, 2020), <https://doi.org/10.1155/2020/6804290>; Rabia Tahir, *A Study on Malware and Malware Detection Techniques*, 2 INT’L J. OF MGMT. ENG’G., 20, 20 (Mar. 8, 2018), <https://www.mecs-press.net/ijeme/ijeme-v8-n2/IJEME-V8-N2-3.pdf>; Mohd Faizal Ab Razak et al., *The Rise of “Malware”*: *Bibliometric Analysis of Malware Study*, 75 J. OF NETWORK AND COMPUT. APPLICATIONS, 58, 58 (Nov. 2016), <https://www.sciencedirect.com/science/article/pii/S1084804516301904>.

1 can generate virtually undetectable malware, and at massive scale, to thwart security systems and
2 jeopardize entire governments.

3 155. Malware attacks have sabotaged entire governments before. For example, in 2022,
4 the Russian Conti Group enacted a weeks-long attack on 27 different ministries in the Costa Rican
5 government.¹⁵⁴ The malware deployed was ransomware, a software that encrypts critical
6 information, denying access to its rightful owner and threatening its destruction if payment is not
7 made.¹⁵⁵ Costa Rica’s president declined to pay the \$20M ransom, but a standoff ensued leaving
8 parts of Costa Rica’s digital infrastructure in shambles, disrupting public healthcare and the pay of
9 its workers.¹⁵⁶

10 156. Healthcare providers are also often targeted by malware, and increasingly so. For
11 example, a major software provider for the UK’s National Health System sustained a ransomware
12 attack from an unknown group last summer.¹⁵⁷ The attack had real impact on the health of millions,
13 disrupting ambulance dispatch, appointment scheduling, and emergency prescriptions, among other
14 things.¹⁵⁸ Ransomware attacks on health care providers have doubled from 2016 to 2021, exposing
15 the sensitive health information of 42M individuals.¹⁵⁹

16 157. ***The Products supercharge Malware:*** In 2012, 33% of malware went undetected by
17 antivirus software.¹⁶⁰ In the last decade, malware has become ever more sophisticated, and ever
18 more capable of thwarting detection. But now, with the assistance of the Products, malware can
19 become undetectable in new ways, at scale, because ChatGPT can be used to create “mutating, or
20

21 ¹⁵⁴ Christine Murray & Mehul Srivastava, *How Conti Ransomware Group Crippled Costa Rica-
Then Fell Apart*, FIN. TIMES (July 9, 2022), <https://www.ft.com/content/9895f997-5941-445c-9572-9cef66d130f5>.

22 ¹⁵⁵ *Id.*

23 ¹⁵⁶ *Id.*

24 ¹⁵⁷ Vedere Labs, *Ransomware in Healthcare: The NHS Example and What the Future Holds*, SEC.
BOULEVARD (Aug. 25, 2022), <https://securityboulevard.com/2022/08/ransomware-in-healthcare-the-nhs-example-and-what-the-future-holds/>.

25 ¹⁵⁸ *Id.*

26 ¹⁵⁹ Hannah T. Neprash et al., *Trends in Ransomware Attacks on US Hospitals, Clinics, and Other
Health Care Delivery Organizations, 2016–2021*, JAMA HEALTH FORUM (Dec. 29, 2022),
27 <https://jamanetwork.com/journals/jama-health-forum/fullarticle/2799961>.

28 ¹⁶⁰ Markus Kammerstetter et al., *Vanity, Cracks, and Malware: Insights into the Anti-Copy
Protection Ecosystem*, ASS’N. FOR COMPUTING MACHINERY 809, 818 (Oct. 16, 2012),
<https://doi.org/10.1145/2382196.2382282>.

1 polymorphic” malware.¹⁶¹ Polymorphic malware has a mutation engine with self-propagating code
2 that allows it to rapidly change its appearance and composition.¹⁶² This malware can change its
3 entire make-up, so that malware detectors, reactionary by nature, will not recognize its newer,
4 ongoing permutations.¹⁶³

5 158. ChatGPT can build the requisite polymorphic code, using its API at runtime to deploy
6 advanced malware attacks that evade detection by security systems designed to thwart malware,
7 such as endpoint detection and response (EDR) applications.¹⁶⁴ Recently, researchers designed a
8 simple, executable file that corresponds with ChatGPT’s API in real time “to generate dynamic,
9 mutating versions of malicious code,” making it extremely difficult to detect using existing
10 cybersecurity tools.¹⁶⁵

11 159. While the most recent iterations of ChatGPT purport to “disallow” potential prompt
12 injections for generating polymorphic malware, this supposed guardrail for safety is woefully
13 inadequate: cleverly worded inputs, used by developers of malware, easily circumvent ChatGPT’s
14 content filters with a practice commonly referred to as “prompt engineering.”¹⁶⁶

15 160. Thus, Mackenzie Jackson, developer advocate at cybersecurity company GitGuardian
16 warns that, as generative models become more advanced, “AI may end up creating malware that
17 can only be detected by other AI systems for defense. What side will win at this game is anyone’s
18 guess.”¹⁶⁷ To knowingly put this enhanced ability to sabotage governments, health care systems,
19 and any other number of targets into the hands of everyday people worldwide without adequate
20 safeguards is emblematic of Defendants’ gross negligence and underscores the need for immediate
21 judicial intervention.

24 ¹⁶¹ Shweta Sharma, *ChatGPT Creates Mutating Malware That Evades Detection by EDR*, CSO
25 ONLINE (June 6, 2023, 1:59 PM), [https://www.csoonline.com/article/3698516/chatgpt-creates-
mutating-malware-that-evades-detection-by-edr.html](https://www.csoonline.com/article/3698516/chatgpt-creates-mutating-malware-that-evades-detection-by-edr.html).

26 ¹⁶² *Id.*

27 ¹⁶³ *Id.*

28 ¹⁶⁴ *Id.*

¹⁶⁵ *Id.*

¹⁶⁶ *Id.*

¹⁶⁷ *Id.*

1 **4. Autonomous Weapons**

2 161. AI also poses a unique threat to international security and human rights through the
3 development of autonomous weapons known as “Slaughterbots,” otherwise known as “lethal
4 autonomous weapons systems” or “killer robots,” which are weapons systems that use AI to
5 identify, select, and kill human targets without intervention.¹⁶⁸ As one humanitarian organization
6 explained, “[w]eapons that use algorithms to kill, rather than human judgment, are immoral and a
7 grave threat to national and global security.”¹⁶⁹

8 162. The risk that unregulated AI like the Products pose via autonomous weapons is “not
9 a far-fetched danger for the future, but a clear and present danger.”¹⁷⁰ Such weapons have already
10 nearly killed a foreign head of state, and due to the rapid commercial proliferation of open-source
11 AI, “could be built today by an experienced hobbyist for less than \$1,000.”¹⁷¹

12 163. Defendants’ conduct exacerbates the problem. There is already an early autonomous
13 implementation of ChatGPT known as “Chaos GPT” which is being touted as “empowering GPT
14 with Internet and Memory to Destroy Humanity.”¹⁷² Chaos-GPT is predicated on an open source
15 application that uses Defendants’ GPT-4, and was designed by an anonymous user to be a
16 “destructive, power-hungry, manipulative AI.”¹⁷³ With only those parameters set by the user,
17 Chaos-GPT returned a list of objectives it set for itself. One was to “destroy humanity.” Another
18 was to “cause chaos and destruction” by creating “widespread suffering.”¹⁷⁴ Next, Chaos-GPT, of
19 its own “volition,” prepared a plan in support of these objectives – and then it searched the internet
20 for weapons of mass destruction seeking to obtain one.¹⁷⁵

21 _____
22 ¹⁶⁸ See *Slaughterbots Are Here*, AUTONOMOUS WEAPONS (Feb. 23, 2023),
23 <https://autonomousweapons.org/> (discussing Latin American and the Caribbean Conference on the
24 Social and Humanitarian Impact of Autonomous Weapons).

25 ¹⁶⁹ *Id.*

26 ¹⁷⁰ Kai-Fu Lee, *The Third Revolution in Warfare*, THE ATLANTIC (Sept. 11, 2021),
27 [https://www.theatlantic.com/technology/archive/2021/09/i-weapons-are-third-revolution-
28 warfare/620013/](https://www.theatlantic.com/technology/archive/2021/09/i-weapons-are-third-revolution-warfare/620013/).

¹⁷¹ *Id.*

¹⁷² Jose Antonio Lanz, *Meet ChaosGPT: An AI Tool That Seeks to Destroy Humanity*, DECRYPT
(Apr. 13, 2023), <https://decrypt.co/126122/meet-chaos-gpt-ai-tool-destroy-humanity>.

¹⁷³ *Id.*

¹⁷⁴ *Id.*

¹⁷⁵ *Id.*

1 164. Experts warn that advancements in AI like those accomplished by the Products, “will
2 accelerate the near-term future of autonomous weapons.”¹⁷⁶ While it is believed artificial
3 intelligence at a level equal to or higher than human intelligence is a prerequisite to truly
4 autonomous weaponry, the unfettered commercial deployment of the Products naturally escalates
5 this risk as their widespread use continually “enhances” the AI’s capabilities – and without sufficient
6 moral or ethical guardrails, as sought in this Action.

7 **C. Opportunity on the Other Side**

8 165. While leading experts agree on the grave risks posed by the Products, and the need
9 for a temporary pause in their commercial deployment, it is important to understand the full picture
10 of why this Action matters. It is not just to contain the risks to society and harms happening right
11 now, including the supercharged spread of disinformation, the obliteration between truth and fiction,
12 deepfakes designed to harass, harm, and commit fraud, and more. It is not just to halt Defendants’
13 ongoing disregard for the privacy and property interests of millions, and to remedy those violations.
14 It is not just to avoid the collapse of civilization as we know it and as Mr. Altman himself recognizes
15 is possible.¹⁷⁷ Naturally, all of these things warrant the comparatively measured relief Plaintiffs and
16 the Classes seek. But beyond all of this, the Action matters to ensure humankind can *realize the*
17 *tremendous opportunity for advancement and prosperity* that awaits us, on the other side of a
18 commercial pause.

19 166. By pausing now, “[h]umanity can enjoy a flourishing future.”¹⁷⁸ It will enable the
20 joint development and implementation of shared safety protocols, overseen by independent outside
21 experts, to manage the risks and render the Products safe to usher in an exciting new era of progress
22 for all. For example, with adequate safeguards, the Products will be positioned to revolutionize
23 healthcare for good, by helping to discover new drugs to save lives and potentially find cures for
24

25 ¹⁷⁶ Lee, *supra* note 170.

26 ¹⁷⁷ David Meyer, *Sam Altman Has Signed a New Open Letter on A.I.’s Dangers: Here’s What’s*
27 *Different About This ‘Extinction’ Statement*, FORTUNE MAG. (May 30, 2023, 9:55 AM),
[https://fortune.com/2023/05/30/sam-altman-has-signed-a-new-open-letter-on-a-i-s-dangers-heres-](https://fortune.com/2023/05/30/sam-altman-has-signed-a-new-open-letter-on-a-i-s-dangers-heres-whats-different-about-this-extinction-statement/)
28 [whats-different-about-this-extinction-statement/](https://fortune.com/2023/05/30/sam-altman-has-signed-a-new-open-letter-on-a-i-s-dangers-heres-whats-different-about-this-extinction-statement/).

¹⁷⁸ *Pause Giant AI Experiments: An Open Letter*, FUTURE OF LIFE INST. (Mar. 22, 2023),
<https://futureoflife.org/open-letter/pause-giant-ai-experiments/>.

1 cancer and other deadly diseases. With adequate safeguards, the Products can contribute not only to
2 our everyday efficiency, artistic expression, joy and more, but also to the greater societal good by
3 advancing human rights, promoting social justice, reducing inequities, and empowering
4 marginalized groups.

5 167. With adequate safeguards, including a moral and ethical code, the Products can help
6 detect and prevent human rights violations rather than cause them; they can help combat human
7 discrimination and bias rather than replicate, encourage, and exacerbate humankind's worst
8 impulses.¹⁷⁹ On the other side of the pause, the Products can responsibly foster global cooperation,
9 collaboration, and peace by facilitating communication, learning, and understanding across cultures
10 and languages rather than starting world wars with disinformation and the unchecked capacity for
11 autonomous weaponry. Likewise, the Products can aid in the ongoing search for truth, by enabling
12 breakthroughs in math, science, and more, that humans might never alone make, rather than forever
13 obliterating the line between truth and fiction altogether.

14 168. We can have this AI, the one that enriches our lives, that works for people, and that
15 works for human benefit, that is "helping us cure cancer, that is helping us find climate solutions,"
16 but leading experts agree, not without a pause on the Products' unchecked commercial proliferation:
17 "[W]hen we're in an arms race to deploy AI to every human being on the planet as fast as possible
18 with as little testing as possible, that's not an equation that's going to end well."¹⁸⁰ The current
19 scenario stands only to enrich Defendants, while destabilizing the world.

20 169. Defendants have released Products to the entire world, that they know and readily
21 recognize could someday result in societal collapse; that even they, the creators, cannot fully
22 understand, predict, or reliably control; thus, any attempt now by Defendants to politicize this
23

24 ¹⁷⁹ See generally Cade Metz and Karen Weise, *A Tech Race Begins as Microsoft Adds A.I. to Its*
25 *Search Engine*, THE N.Y. TIMES (Feb. 7, 2023),
26 <https://www.nytimes.com/2023/02/07/technology/microsoft-ai-chatgpt-bing.html> ("The new
27 chatbots do come with baggage. They often do not distinguish between fact and fiction. They can
28 generate language that is biased against women and people of color. And experts worry that
people will use them to spread lies at a speed they could not in the past.").

¹⁸⁰ Jason Abbruzzese, *The Tech Watchdog That Raised Alarms About Social Media Is Warning*
About AI, NBC NEWS (Mar. 22, 2023), [https://www.nbcnews.com/tech/tech-news/tech-watchdog-
raised-alarms-social-media-warning-ai-rcna76167](https://www.nbcnews.com/tech/tech-news/tech-watchdog-
raised-alarms-social-media-warning-ai-rcna76167).

1 action, to attack the class action device or those brave enough to stand up to corporate greed and
2 irresponsibility of this magnitude at this pivotal moment in history, will fail. All people of good will
3 on both sides of the aisle and from every background are united and resolute in the need for
4 intervention. That is because we all want to live in a world where technology serves our shared
5 values of freedom, justice, dignity, equality, prosperity, privacy and security, not where Products
6 exist that undermine these ideals.

7 170. In an often divided and polarized world, it is telling how so many have been able to
8 unite around these truths: (i) the current state of AI governance is insufficient to address the threats
9 posed by the Products; (ii) the lack of transparency, accountability, oversight, and regulation
10 surrounding the Products and Defendants suddenly deploying them for profit worldwide has
11 resulted in a ticking time bomb in the hands of those motivated to harm the American people; (iii)
12 the gap must be closed between the rapid pace of the Products' development on the backs of stolen
13 personal data on the one hand, and the slow progress of AI policy on the other; and (iv) a temporary
14 pause on the commercial deployment of the Products is necessary and justified to prevent
15 irreversible damage to humanity and society.

16 171. Critically, the injunctive relief sought in this Action seeks only to pause the unfettered
17 and further commercial deployment of the Products, with AI research and development otherwise
18 continuing unaffected. That is because of an equally important truth on which all agree: the United
19 States must remain aggressively locked into the worldwide AI arms-race, set off by Defendants'
20 launch of the Products (for better or worse), to ensure this powerful technology is developed and
21 deployed for good around the world, and to block the potential harms from those world powers
22 currently leveraging AI like the Products to build technological weapons as powerful as the nuclear
23 bomb. Thus, the only "setback" here will be to Defendants' corporate bank accounts, while the
24 American people stand to (re)gain their fundamental right to privacy as well as just compensation
25 for the mass theft of personal data on which Defendants built and continue to run the Products.

26
27
28

1 **III. DEFENDANTS’ CONDUCT VIOLATES ESTABLISHED PROPERTY AND**
2 **PRIVACY RIGHTS**

3 **A. Defendants’ Web-Scraping Theft**

4 172. Defendants’ first category of theft and misappropriation stems from their secret
5 scraping of the internet. This violated both the property rights and privacy rights of all individuals
6 whose personal information was scraped and then incorporated through misappropriation into
7 Defendants’ Products.

8 173. Defendants’ initial web scraping was done largely in secret, without the consent of
9 any individuals whose personal and identifying information was scraped, much less all of the
10 website operators themselves. This violated not only the Terms of Use of various websites but also
11 the rights of each and every individual to opt out of such collection under California and other state
12 and federal laws. Without any notice to the public, no one can be said to have consented to the
13 collection of their online personal data, history, web practices and other personal and identifying
14 information.

15 174. By the time the public learned of Defendants’ web scraping practices in late Fall of
16 2022, when ChatGPT was released, it was too late to meaningfully exercise their privacy rights
17 outside of this lawsuit — their internet history had been scraped, consumed, and integrated into the
18 large language models from which the Products were born.

19 175. While Defendants’ massive theft of personal information at scale is unmatched in
20 history, it is reminiscent of the Clearview AI scandal in 2020. Clearview is a company that uses
21 facial recognition technology to identify individuals based on their online photos.¹⁸¹ To create its
22 product, Clearview scraped billions of publicly available photos from various websites and social
23 media platforms.¹⁸² As with Defendants, this illegal scraping was done without the consent of users
24 or the website owners themselves, and without registering as a data broker under California or
25

26 ¹⁸¹ Tate Ryan-Mosley, *The NYPD Used a Controversial Facial Recognition Tool. Here’s What*
27 *You Need to Know*. MIT TECH. REV., (Apr. 9, 2021),
28 www.technologyreview.com/2021/04/09/1022240/clearview-ai-nypd-emails/.

¹⁸² Will Knight, *Clearview AI Has New Tools to Identify You in Photos*, WIRED (Oct. 4, 2021),
<https://www.wired.com/story/clearview-ai-new-tools-identify-you-photos/>.

1 Vermont Law.¹⁸³

2 176. Just like Defendants, Clearview used the stolen information to build its AI product.¹⁸⁴
 3 Clearview then sold access to the product to law enforcement agencies, private companies, and other
 4 governmental agencies.¹⁸⁵ Defendants' business model is the same: scrape information off the
 5 internet, in secret without any notice and consent in violation of the law, use it to build AI products,
 6 and then sell access to the Products for commercial gain.

7 177. Clearview's illegal scraping practices also went undetected for years, until it was laid
 8 bare by a New York Times expose.¹⁸⁶ The public was rightfully upset, as were state and federal
 9 regulators. The Vermont Attorney General sued Clearview in March 2020 for violating data broker
 10 and consumer protection laws, alleging that Clearview fraudulently acquired brokered personal
 11 information through its scraping practices and exposed consumers to various risks and harms.¹⁸⁷
 12 Clearview was also sued by several individuals and organizations in California and elsewhere.¹⁸⁸

13 178. As a result of these lawsuits and public scrutiny, Clearview ultimately registered as a
 14 data broker in both California and Vermont. Although Defendants employ the same business model
 15 as Clearview, they have failed to register as data brokers under applicable law. By failing to do so
 16 prior to scraping the internet, Defendants violated the rights of millions. Plaintiffs and the Classes
 17 had a right to know what personal information Defendants were scraping and collecting and how it
 18 would be used, a right to delete their personal information collected by Defendants, and a right to

19 _____
 20 ¹⁸³ Robert Hart, *Clearview AI Fined \$9.4 Million in UK for Illegal Facial Recognition Database*,
 21 FORBES (May 23, 2022), <https://www.forbes.com/sites/roberthart/2022/05/23/clearview-ai-fined-94-million-in-uk-for-illegal-facial-recognition-database/?sh=73d5a0f71963>.

21 ¹⁸⁴ *Id.*

22 ¹⁸⁵ Drew Harwell, *Clearview AI to Stop Selling Facial Recognition Tool to Private Firms*, THE
 23 WASH. POST (May 9, 2022), <https://www.washingtonpost.com/technology/2022/05/09/clearview-illinois-court-settlement/>.

24 ¹⁸⁶ Dave Gershgor, *Is There Any Way Out of Clearview's Facial Recognition Database?*, THE
 25 VERGE (June 9, 2021), <https://www.theverge.com/22522486/clearview-ai-facial-recognition-avoid-escape-privacy>.

26 ¹⁸⁷ *Attorney General Donovan Sues Clearview AI for Violations of Consumer Protection Act and
 27 Data Broker Law*, OFF. OF VT. ATT'Y GEN. (Mar. 10, 2020),
 28 <https://ago.vermont.gov/blog/2020/03/10/attorney-general-donovan-sues-clearview-ai-violations-consumer-protection-act-and-data-broker-law>.

¹⁸⁸ Johana Bhuiyan, *Clearview AI Uses Your Online Photos to Instantly ID You. That's A Problem, Lawsuit Says*, L.A. TIMES (Mar. 9, 2021),
<https://www.latimes.com/business/technology/story/2021-03-09/clearview-ai-lawsuit-privacy-violations>.

1 opt out of the use of that information to build the Products.

2 179. Defendants' violation of the law is ongoing as they continue to collect personal
3 brokered information by scraping the internet without registering as data brokers or otherwise
4 providing notice or seeking consent from anyone. Plaintiffs and the Classes have a right to opt out
5 of this ongoing scraping of internet information but no mechanism to exercise that right, absent the
6 injunctive relief sought in this Action.

7 **B. Defendants' Web Scraping Violated Plaintiff's Property Interests**

8 180. Courts recognize that internet users have a property interest in their personal
9 information and data. *See Calhoun v. Google, LLC*, 526 F. Supp. 3d 605, 635 (N.D. Cal. 2021)
10 (recognizing property interest in personal information and rejecting Google's argument that "the
11 personal information that Google allegedly stole is not property"); *In re Experian Data Breach*
12 *Litigation*, SACV 15-1592 AG (DFMx), 2016 U.S. Dist. LEXIS 184500, at *14 (C.D. Cal. Dec. 29,
13 2016) (loss of value of personal identifying information is a viable damages theory); *In re Marriott*
14 *Int'l Inc. Customer Data Sec. Breach Litig.*, 440 F. Supp. 3d 447, 460-61 (D. Md. 2020) ("The
15 growing trend across courts that have considered this issue is to recognize the lost property value of
16 this [personal] information."); *Simona Opris v. Sincera*, No. 21-3072, 2022 U.S. Dist. LEXIS
17 94192, at *20 (E.D. Pa. May 23, 2022) (collecting cases).

18 181. Plaintiffs' and Class Members' property rights in the personal data and information
19 that they have generated, created, or provided through various online platforms thus includes the
20 right to possess, use, profit, sell, and exclude others from accessing or exploiting that information
21 without consent or remuneration.

22 182. The economic value of this property interest in personal information is well
23 understood, as a robust market for such data drives the entire technology economy. As experts have
24 noted, the world's most valuable resource is "no longer oil, but data," and has been for years now.¹⁸⁹

25 183. A single internet user's information can be valued anywhere from \$15 to \$40, and
26

27 _____
28 ¹⁸⁹ *The World's Most Valuable Resource Is No Longer Oil, but Data*, THE ECONOMIST (May 6,
2017), <https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data>.

1 even more.¹⁹⁰ Another study found that an individual’s online identity can be sold for \$1,200 on the
 2 dark web.¹⁹¹ Defendants’ misappropriation of every piece of data available on the internet, and with
 3 it, millions of internet users’ personal information without consent, thus represents theft of a value
 4 unprecedented in the modern era of technology.

5 184. Writing for the Harvard Law Review, Professor Paul M. Schwartz underscored the
 6 value of personal data, as follows: “Personal information is an important currency in the new
 7 millennium. The monetary value of personal data is *large* and still *growing*, [and that’s why]
 8 corporate America is moving quickly to profit from the trend.”¹⁹² The data forms a critical “corporate
 9 asset.”

10 185. Other experts concur: “[S]uch vast amounts of collected data have obvious and
 11 substantial economic value. Individuals’ traits and attributes (such as a person’s age, address,
 12 gender, income, preferences... [their] clickthroughs, comments posted online, photos updated to
 13 social media, and so forth) are increasingly regarded as business assets[.]”¹⁹³

14 186. Because personal data is valuable personal property, market exchanges now exist
 15 where internet users like Plaintiffs and putative class members can sell or monetize their own
 16 personal data and internet usage information.¹⁹⁴ For example, Facebook has offered to *pay* users for
 17 their voice recordings.¹⁹⁵ By contrast and as alleged herein upon information and belief, Defendants

18
 19 ¹⁹⁰ *Id.*

20 ¹⁹¹ Maria LaMagna, *The Sad Truth About How Much Your Facebook Data is Worth on the Dark*
 21 *Web*, MARKETWATCH (June 6, 2018), [https://www.marketwatch.com/story/spooked-by-the-
 facebook-privacy-violations-this-is-how-much-your-personal-data-is-worth-on-the-dark-web-
 2018-03-20](https://www.marketwatch.com/story/spooked-by-the-facebook-privacy-violations-this-is-how-much-your-personal-data-is-worth-on-the-dark-web-2018-03-20).

22 ¹⁹² Paul M. Schwartz, Property, Privacy, and Personal Data, 117 HARV. L. REV. 2056, 2056 (May,
 2004).

23 ¹⁹³ Alessandro Acquisti et al., *The Economics of Privacy*, 54(2) J. OF ECON. LITERATURE 442, 444
 (Mar. 8, 2016).

24 ¹⁹⁴ Kevin Mercandante, *10 Apps for Selling Your Data for Cash*, BEST WALLET HACKS,
 25 <https://wallethacks.com/apps-for-selling-your-data/> (last updated Apr. 20, 2023); Kari Paul,
 26 *Facebook Launches Apps That Will Pay Users for Their Data*, THE GUARDIAN (June 11, 2019)
 27 <https://www.theguardian.com/technology/2019/jun/11/facebook-user-data-app-privacy-study>;
 Saheli Roy Choudry & Ryan Browne, *Facebook Pays Teens to Install an App That Could Collect*
 28 *All Kinds of Data*, CNBC (Jan. 29, 2019), [https://www.cnbc.com/2019/01/29/facebook-paying-
 users-to-install-app-to-collect-data-techcrunch.html](https://www.cnbc.com/2019/01/29/facebook-paying-users-to-install-app-to-collect-data-techcrunch.html).

¹⁹⁵ Tim Bradshaw, *Facebook Offers to Pay Users for Their Voice Recordings*, FIN. TIMES (Feb.
 21, 2020), <https://www.ft.com/content/42f6b93c-54a4-11ea-8841-482eed0038b1>.

1 simply *took* millions of text files, voice recordings, and facial scans from across the internet —
2 without any consent from putative class members, much less personal remuneration to them. Theft
3 of this nature is not only unprecedented and unjust, but also dangerous. As noted in Section II, it
4 puts millions at risk for their likeness to be cloned to perpetrate fraud, or to embarrass or otherwise
5 harm them.

6 187. Moreover, the law specifically recognizes a legal interest in unjustly earned profits
7 based on unauthorized harvesting of personal data, and “this stake in unjustly earned profits exists
8 regardless of whether an individual planned to sell his or her data or whether the individual’s data
9 is made less valuable.”¹⁹⁶

10 188. Defendants have been unjustly enriched by their theft of personal information as its
11 billion-dollar AI business, including ChatGPT and beyond, was built on harvesting and monetizing
12 Internet users’ personal data. Thus, Plaintiffs and the Classes have a right to disgorgement and/or
13 restitution damages representing the value of the stolen data and/or their share of the profits
14 Defendants earned thereon.

15 **C. Defendants’ Web Scraping Violated Plaintiffs’ Privacy Interests**

16 189. In addition to property rights, internet users maintain privacy interests in personal
17 information even if it is posted online, and experts agree the collection, processing, and further
18 dissemination of this information can create distinct privacy harms.¹⁹⁷

19 190. For example, the aggregation of collected information “can reveal new facts about a
20 person that she did not expect would be known about her when the original, isolated data was
21 collected.”¹⁹⁸ Even a small subset of “public” private information can be used to harm the privacy
22 interests of internet users. One example is when researchers analyzed public tweets to identify users
23 with mental health issues; naturally, Twitter users did not consent or expect their data to be used in
24 that way, to potentially reveal new, highly personal information about them.¹⁹⁹ If that analysis were
25 made public, or used commercially, that would pose significant and legally cognizable privacy

26 ¹⁹⁶ *In re Facebook, Inc. Internet Tracking Litigation*, 956 F.3d 589, 600 (9th Cir. 2020).

27 ¹⁹⁷ Geoffrey Xiao, *Bad Bots: Regulating the Scraping of Public Information*, 34(2) HARV. J.L. &
TECH., 701, 706, 732 (2021).

28 ¹⁹⁸ Daniel J. Solove, *A Taxonomy of Privacy*, 154 U. PA. L. REV. 477, 493 (2006).

¹⁹⁹ Xiao, *supra* note 197, at 707.

1 harms.

2 191. Another reason users retain privacy interests in their personal data on the internet,
3 even when it is technically ‘public,’ is the reasonable expectation of “obscurity” i.e., “the notion
4 that when our activities or information is unlikely to be found, seen, or remembered, it is, to some
5 degree safe.”²⁰⁰ Privacy experts note users’ reasonable expectation that most of the Internet will
6 simply ignore their individual posts. Moreover, “[t]he passage of time also makes information
7 obscure: no one remembers your MySpace pictures from fifteen years ago.”²⁰¹

8 192. Internet users’ reasonable expectations are also informed by the known transaction
9 costs that, typically, would “prevent[] someone from collecting all your photos from every social
10 media site you have ever used – ‘just because information is hypothetically available does not mean
11 most (or even a few) people have the knowledge and ability to access [‘public’ private]
12 information.”²⁰²

13 193. When users post information on the internet, “they do so believing that their
14 information will be obscure and in an environment of trust” on whichever site they post. Users
15 expect a level of privacy— they “**do not expect their information to be swept up by data**
16 **scraping.**” Thus, according to experts, the privacy problem with “widescale, automated collection
17 of personal information via scraping,” is that it “destroys” reasonable user expectations including
18 the right to “obscurity” by reducing the typical transaction costs and difficulties in accessing,
19 collecting, and understanding personal information at scale.²⁰³

20 194. Scraping therefore illegally enables the use of personal information in ways which
21 reasonable users could not have anticipated. In respect of Defendants’ surreptitious scraping at
22 unprecedented scale, it means all items users have posted on the internet have now been collected,
23 including their voice recordings and images – arming Defendants with the ability to create a digital
24 clone of each internet user to anticipate and manipulate their next move. Plaintiffs and the Classes
25 did not consent to such use of their personal information. As privacy experts note, “even if a user
26

27 ²⁰⁰ Woodrow Hartzog, *The Public Information Fallacy*, 99 BOS. L. REV. 459, 515 (2019).

28 ²⁰¹ Xiao, *supra* note 197, at 708-09.

²⁰² *Id.* at 709.

²⁰³ *Id.*

1 makes the affirmative choice to make [an internet post public], she manifests an intent to participate
2 in an obscure and trustworthy environment, not an intent to participate in data harvesting.”²⁰⁴

3 195. Worse, Plaintiffs and the Classes could not have known Defendants were collecting
4 their personal information, because Defendants did it without notice to anyone, in violation of
5 California law which required them to register with the state as data brokers.²⁰⁵

6 196. Introducing these data broker laws, the California assembly stated its intent:
7 “[C]onsumers are generally not aware that data brokers possess their personal information, how to
8 exercise their right to opt out, and whether they can have their information deleted, as provided by
9 California law.” Thus, “it is the intent of the Legislature to further Californians’ right to privacy by
10 giving consumers an additional tool to help control the collection and sale of their personal
11 information by requiring data brokers to register annually with the Attorney General and provide
12 information about how consumers may opt out of the sale of their personal information.”²⁰⁶

13 197. “Sale” of information includes “making it available” to others for consideration,
14 which Defendants have done by commercializing the stolen data into ChatGPT and building a
15 billion-dollar business from it. Despite scraping information for this express purpose, Defendant
16 OpenAI did not, and still has not, registered with the State of California as required.

17 198. Experts acknowledge the “serious privacy harms” inherent in the type of entirely
18 “covert information” collection in which Defendants engaged.²⁰⁷ It “undermines individual
19 autonomy and free choice.”²⁰⁸ The lack of notice, including under California’s data broker laws,
20 “excludes individuals from the data collection process, making individuals feel powerless in
21 controlling how their data is used.”²⁰⁹ This is not just a feeling—as described *supra*, the harm is
22 concrete economic injury given the robust market for personal information.

23 199. Without notice of Defendants’ scraping practices, users were also denied the ability
24

25 ²⁰⁴ *Id.* at 711.

26 ²⁰⁵ Cal. Civ. Code § 1798.99.80(d).

27 ²⁰⁶ Assemb. B. 1202, 2019-2020 Reg. Sess. (Cal. 2019) (as discussed in Xiao, *supra* note 197, at
714-715).

28 ²⁰⁷ Xiao, *supra* note 197, at 719.

²⁰⁸ *Id.*

²⁰⁹ *Id.*

1 to engage in self-help, by choosing to make obscure but technically publicly-available information
 2 private – and the lack of notice precluded users from exercising their statutory data privacy rights,
 3 such as the right to request deletion.²¹⁰ Instead, Plaintiffs’ and the Classes’ internet histories are now
 4 embedded in Defendants’ AI products with no recourse other than the damages and injunctive relief
 5 requested in this Action.

6 **D. Defendants’ Business Practices are Offensive to Reasonable People and Ignore**
 7 **Increasingly Clear Warnings from Regulators**

8 200. Defendants’ mass scraping of personal data for commercialization has sparked
 9 outrage over the legal and privacy implications of Defendants’ practices. Those aware of the full
 10 extent of the misappropriation are fearful and anxious about how Defendants used their “digital
 11 footprint” and about how Defendants might use all that personal information going forward. Absent
 12 the relief sought in this Action, there will be no limits on such future use. The public is also
 13 concerned about how all of their personal information might be accessed, shared, and misused *by*
 14 *others*, now that it is forever embedded into the large language models on which the Products run.

15 201. The outrage makes sense: Defendants admit the Products might evolve to act against
 16 human interests, and that regardless, they are unpredictable. Thus, by collecting previously obscure
 17 and personal data of millions and permanently entangling it with the Products, Defendants
 18 knowingly put Plaintiffs and the Classes in a zone of risk that is *incalculable* — but unacceptable
 19 by any measure of responsible data protection and use.

20 202. The extent to which Defendants stand to profit from the unprecedented privacy risks
 21 they were willing to take—with data that is not theirs—is especially offensive to everyday people.
 22 As one explained, “Using AI as it stands right now is *normalizing the illegal mass scraping* of
 23 everyone’s data regardless of their nature, just to make the top even richer and forfeit any means we
 24 have to protect our work *and who we are as humans*. This should not be encouraged and
 25 tolerated.”²¹¹ The outrage stems, in part, from this uncontestable truth: “None of this would have
 26

27 _____
²¹⁰ *Id.* at 720.

28 ²¹¹ @coffeeseed, TWITTER (May 11, 2023),
<https://twitter.com/CoffeeSeed/status/1656634134616211461>.

1 been possible without data – *our data* – collected and used without our permission.”²¹²

2 203. In this new era of AI, we cannot allow widescale illegal data scraping to become a
3 commercial norm; otherwise, privacy as a fundamental right will be relegated to the dustbin of
4 history. Underscoring the need for court intervention, AI researcher Remmelt Ellen remarked
5 simply, “[i]llegal scraping needs to be addressed.”²¹³

6 204. The public is also troubled by the lack of just compensation for the use of their
7 personal data. One AI large language model developer stated it plainly: “If your data is used,
8 companies should cough up.”²¹⁴ Otherwise, according to a more complete critique of the current
9 business model, AI is just “pure primitive accumulation”—taking from the masses to enrich a few,
10 i.e., Silicon Valley tech companies and their billionaire owners.²¹⁵

11 205. While the past, and ongoing, misappropriation of valuable personal information is bad
12 enough, the Products also stand to altogether eliminate future income for millions, due to the
13 widespread unemployment they are expected to cause over time. No one has consented to the use
14 of their personal information to build this destabilized future of social unrest and worsening poverty
15 for everyday people, while the pockets of OpenAI and Microsoft are lined with profit.

16 206. As OpenAI itself once acknowledged, albeit when still purely not-for-profit, the
17 Company would need to fund a universal basic income (UBI) if the Products were ever developed
18 and deployed for widespread public use, because they would eliminate so many jobs. Even now,
19 Mr. Altman’s “grand idea is that OpenAI will capture much of the world’s wealth through the
20 creation of A.G.I. and then redistribute this wealth to the people.”²¹⁶ Given Defendants’ sudden
21 deployment of the Products across virtually every industry using data that was not theirs, this future
22 should begin now, with legal or equitable redistribution of Defendants’ ill-gotten gains. Others have

23
24 ²¹² Gal, *supra* note 5.

25 ²¹³ @RemmeltE, TWITTER (Apr. 10, 2023),
<https://twitter.com/RemmeltE/status/1645499008075407364>.

26 ²¹⁴ @yudhanjaya, TWITTER (June 9, 2023),
<https://twitter.com/yudhanjaya/status/1667391709679095808>.

27 ²¹⁵ Bridle, *supra* note 76.

28 ²¹⁶ Cade Metz, *The ChatGPT King Isn’t Worried, but He Knows You Might Be*, THE N.Y. TIMES
(Mar. 31, 2023), [https://www.nytimes.com/2023/03/31/technology/sam-altman-open-ai-
chatgpt.html](https://www.nytimes.com/2023/03/31/technology/sam-altman-open-ai-chatgpt.html).

1 noted that a portion of the profits generated by Defendants can be funneled back “to everyone who
2 contributed content.” This would include “basically everyone,” given the scope of the initial and
3 ongoing theft of personal information by Defendants.²¹⁷

4 207. To avoid the unjust enrichment of Defendants, this Court sitting in equity has the
5 power to order a “data dividend” to consumers for as long as the Products generate revenue fueled
6 on the misappropriated data. At the very least, Plaintiffs and the Classes should be personally and
7 directly compensated for the fair market value of their contributions to the large language models
8 on which the Products were built and thrive, in an amount to be determined by expert testimony.
9 Fundamental principles of property law demand such compensation, and everyday people
10 reasonably support it.²¹⁸

11 208. While the property and privacy rights this Action seeks to vindicate are settled as a
12 general matter, their application to business practices surrounding the large language models fueling
13 AI products has not been widely tested under the law. However, just weeks ago, the FTC settled an
14 action against Amazon, in connection with the company’s illegal use of voice data to train the
15 algorithms on which its popular Alexa product runs. That action raised many of the same type of
16 violations alleged in this Action.

17 209. Announcing settlement of the action, the FTC gave a stern public warning to
18 companies like Defendants: “Amazon is not alone in apparently seeking to amass data to refine its
19 machine learning models; right now, with the advent of large language models, the tech industry as
20 a whole is *sprinting* to do the same.”²¹⁹ The settlement, it continued, was to be a message to all:
21 “Machine learning is *no excuse to break the law*... The data you use to improve your algorithms
22 must be *lawfully collected* and *lawfully retained*. Companies would do well to heed this lesson.”²²⁰

23 210. The FTC’s warning comports with FTC Commissioner Rebecca Slaughter’s earlier

24 ²¹⁷ *Id.*

25 ²¹⁸ See e.g., @ianfinlay2000, *Time to Get Paid For Our Data?*, REDDIT (2021),
26 https://www.reddit.com/r/Futurology/comments/qknz3u/time_to_get_paid_for_our_data/ (“[T]he
27 companies are basically stealing our data bc no one knows that they should be getting paid for it”).

28 ²¹⁹ Devin Coldewey, *Amazon Settles with FTC for \$25M After ‘Flouting’ Kids’ Privacy and
Deletion Requests*, TECHCRUNCH (May 31, 2023), <https://techcrunch.com/2023/05/31/amazon-settles-with-ftc-for-25m-after-flouting-kids-privacy-and-deletion-requests/> (emphasis added).

²²⁰ *Id.* (emphasis added).

1 warning, in 2021, in the Yale Journal of Law and Technology.²²¹ Discussing the FTC’s new practice
 2 of ordering “algorithmic destruction,” Commissioner Slaughter explained that “the premise is
 3 simple: when companies collect data illegally, they should not be able to profit from either the data
 4 or any algorithm developed using it.”²²² Commissioner Slaughter believed this enforcement
 5 approach would “send a clear message to companies engaging in illicit data collection in order to
 6 train AI models: *Not worth it.*”²²³ Unfortunately for the millions of consumers impacted by
 7 Defendants’ mass theft of data, Defendants did not heed the warning.

8 **E. Defendants’ Theft of User Data in Excess of Reasonable Consent**

9 211. Defendants’ second category of theft stems from their unrestricted harvesting of data
 10 from Users of the Products, including registered Users of the OpenAI website and Users of
 11 Defendants’ API and/or plug-ins.

12 212. Defendants have made much of the fact that they purportedly “want” to comply with
 13 applicable privacy laws and regulations—and will likely oppose this lawsuit by arguing that
 14 registered users of the Products purportedly “consented” to the widespread theft of their personal
 15 information by virtue of using the Products. This argument is disingenuous for multiple reasons.

16 213. *First:* For those consumers who used ChatGPT plug-ins or API, the various sites’ use
 17 policies did not provide anything approaching informed consent that the consumers’ information
 18 and personal data would be used to train Defendants’ LLMs and would thus be incorporated into
 19 generative AI in a manner that would prevent them from reasonably ever removing their data from
 20 Defendants’ for-profit commercial enterprises. Plaintiffs and Class Members had no idea that
 21 Defendants were and are collecting and utilizing their User Data, including the most sensitive
 22 information, when they engage with ChatGPT which seamlessly incorporated artificial intelligence
 23 in the background.

24 214. Plaintiffs fell victim to Defendants’ unlawful collection and sharing of their sensitive
 25 information acquired through their interactions with Defendants’ Products and websites, as well as

26
 27 ²²¹ Rebecca Kelly Slaughter et al., *Algorithms and Economic Justice: A Taxonomy of Harms and a*
Path Forward for the Federal Trade Commission, 23 YALE J. L. & TECH. 1, 39 (Aug. 2021).

28 ²²² *Id.*

²²³ *Id.* (emphasis added).

1 the hundreds or thousands of applications that now use ChatGPT-based plug-ins or API.²²⁴

2 215. In less than 24 hours after Defendants announced the ability to install plug-ins to
3 ChatGPT, many companies immediately jumped on board and started incorporating their websites
4 within the AI plug-in. In exchange, Defendants received yet another wealth of personal data, once
5 again, without the users' and nonusers' consent. ChatGPT is becoming the single app "to rule them
6 all."²²⁵

7 216. Defendants' AI has become the virtual spy,²²⁶ closely monitoring, recording, and
8 training on the personal data, clicks, searches, inputs, and personal information of millions of
9 unsuspecting individuals who may be using an Instacart to purchase grocery items, a telehealth
10 company to make a doctor's appointment, or simply browsing Expedia to make vacation plans.

11 217. *Second:* Even those who registered for OpenAI accounts and interacted with ChatGPT
12 directly did not give effective consent for Defendants to use their data and personal information in
13 the way they currently do.

14 218. For instance, when Plaintiffs logged in to use the ChatGPT, Defendants were tracking
15 and collecting every piece of information entered into the chatbot—including sensitive information
16 such (1) all details entered into the chatbot; (2) account information users enter when signing up;
17 (3) name; (4) contact details; (5) login credentials; (6) emails; (7) payment information; (8)
18 transaction records; (9) identifying data ChatGPT pulls from users' device or browser, like IP
19 addresses and location; (10) social media information; (11) chat log data; (12) usage data; (13)
20 analytics; and (14) cookies. However, Defendants are also tracking the information from other
21 applications in which their AI is already plugged in – Stripe, Microsoft Teams, Bing, Zillow,
22 Expedia, Instacart, etc. – and using each piece of information to train the AI.

23 219. Plaintiffs, and all Class Members, did not consent to such extensive collection of data,
24 and the use of their data for essentially any purpose to benefit Defendants' businesses – including

25 _____
26 ²²⁴ Matt Burgess, *ChatGPT Has a Big Privacy Problem*, WIRED (Apr. 4, 2023),
<https://www.wired.com/story/italy-ban-chatgpt-privacy-gdpr/>.

27 ²²⁵ Better Product, *OpenAI's Master Plan to Turn ChatGPT into an Everything App*, MEDIUM
28 (Mar. 25, 2023), <https://medium.com/@betterproducts/openais-master-plan-to-turn-chatgpt-into-an-everything-app-1270686074f8>.

²²⁶ *Id.*

1 for training purposes of the AI. In fact, Plaintiffs and all Class Members could not consent to
2 Defendants' conduct because they were unaware their sensitive information would be collected and
3 used in this manner in the first place. Thus, Defendants did not obtain *valid enforceable* consent to
4 collect, use, and store Plaintiffs' and Class Members' sensitive information.

5 220. In the near future, Defendants anticipate adding even more powerful features to the
6 omniscient AI, allowing it to also gather data from audio inputs with their yet another AI—Vall-E.
7 Vall-E has already been developed and allows to process three (3) seconds of a human voice, and
8 be able to speak in such voice in perpetuity. Once activated, Defendants' and their AI's access to
9 human voices and audio inputs will jeopardize the users' and nonusers' privacy even further.

10 221. Defendant OpenAI has also deceptively represented to its users that they can request
11 their private information not be used and, if parents discover that a child has used ChatGPT,
12 Defendant will erase the child's data from the system. This is deceptive because by the time the
13 language model has taken in the information and learned from it, that information has already
14 financially benefited Defendants and cannot be removed from the knowledge base of the language
15 model. Moreover, Defendant OpenAI has stated that, notwithstanding a user's requests to opt out
16 of data collection and sharing, it will still retain some information (though what information will be
17 retained is not specified).

18 222. Currently, a ChatGPT user wanting to opt out of the use of their data and chats for
19 model training is instructed that they can simply turn off chat history (which deprives them of using
20 that functionality themselves) and the application will stop using *new* chat content for training
21 purposes.²²⁷ However, Defendants continue to train their models with the user's information – be it
22 from the prior chats or new chats. Moreover, as noted above, it is impossible to know whether any
23 of the previously used data can effectively be deleted, as once the language model is trained using
24 the data, it becomes part of the model. Additionally, the option of opting out of chat history retention
25 doesn't impact OpenAI's ability to use a user's other personal data gathered during the account
26 creation process for Defendants' own purposes. OpenAI's privacy disclosures are intentionally

27 ²²⁷ Johanna C., *How Do I Turn Off Chat History and Model Training?*, OPENAI,
28 <https://help.openai.com/en/articles/7792795-how-do-i-turn-off-chat-history-and-model-training>
(last visited June 27, 2023).

1 vague about this, noting simply that a user can opt out of chat history retention *or* can submit a form
2 to ask OpenAI not to use or share their data. No guidelines are provided regarding whether or when
3 Defendant might decline to honor such a request, nor how long it takes to process.

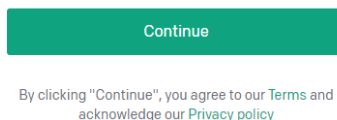
4 223. Furthermore, as commentators have observed, Defendant OpenAI heavily pushes
5 users not to opt out of data collection.²²⁸ Once a user turns off the option for their ChatGPT
6 interactions to be used for training purposes, they are presented constantly with a large green button
7 that encourages them to “Enable chat history.” Nothing on this button notifies users that enabling
8 chat history functionality amounts to reauthorizing OpenAI to save and train Defendants’ models
9 on the user’s data.

10 224. Moreover, it is not clear what information a given user can actually prevent OpenAI
11 from retaining and using in the future, as the company has stated in blog posts that it will retain
12 some data anyway and that some of this data can be used in Defendant OpenAI’s training datasets.²²⁹

13 225. Defendants fail to provide accurate and comprehensive notifications to consumers
14 about the scale of their data sharing practices. Defendants’ admissions within their Privacy Policy
15 do not adequately inform consumers on the breadth of data sharing, resulting in a breach of explicit
16 assurances and a violation of reasonable consumer expectations. By acting in such a manner,
17 Defendants are engaged in data misuse practices that contradict the principles of transparency,
18 accountability, and respect for consumer privacy rights.

19 ***1. OpenAI’s disclosures are not conspicuous.***

20 226. When a consumer attempts to register for an OpenAI account, they are presented with
21 the following image:



22
23
24
25
26 ²²⁸ Natasha Lomas, *How to Ask OpenAI for Your Personal Data to Be Deleted or Not Used to*
Train Its AIs, TECHCRUNCH (May 2, 2023), [https://techcrunch.com/2023/05/02/chatgpt-delete-](https://techcrunch.com/2023/05/02/chatgpt-delete-data/)
27 *data/*.

28 ²²⁹ Yaniv Markovski, *How Your Data Is Used to Improve Model Performance*, OPENAI,
[https://help.openai.com/en/articles/5722486-how-your-data-is-used-to-improve-model-](https://help.openai.com/en/articles/5722486-how-your-data-is-used-to-improve-model-performance)
performance (last visited June 2, 2023).

1 227. When a hyperlink to an agreement is “not conspicuous enough to put [plaintiffs] on
2 inquiry notice,” then the agreement is not binding. *Colgate v. JUUL Labs, Inc.*, 402 F. Supp. 3d
3 728, 764-66 (N.D. Cal. 2019). The Ninth Circuit holds that “even close proximity of the hyperlink
4 to relevant buttons users must click on—without more—is insufficient to give rise to constructive
5 notice.” *Nguyen v. Barnes & Noble Inc.*, 763 F.3d 1171, 1179 (9th Cir. 2014). Instead, courts
6 consider factors such as color, size and font of the hyperlink, and whether the hyperlink is presented
7 alone or in a clutter of text. *See, e.g., Colgate*, 402 F. Supp. 3d at 764; *Selden v. Airbnb, Inc.*, 16-
8 cv-00933 (CRC), 2016 WL 6476934, at *14-15 (D.D.C. Nov. 1, 2016).

9 228. Here, a consumer registering for an OpenAI account is ferried through the process
10 and is provided only small hyperlinks to OpenAI’s Privacy Policy and Terms of Use during the
11 sign-up process. The lettering alerting the potential registrant to the documents is tiny and gray. The
12 consumer need not make any indication that he or she has actually read the documents, nor that they
13 understand the connection between these documents and their creation of an account. Unlike many
14 companies that require a consumer to scroll to the bottom of a privacy policy or other legal
15 document—or at least click a radial purporting to have read the document—an OpenAI registrant
16 need make no affirmative indication that they are aware of the policies whatsoever. As such, there
17 is no binding agreement between Defendant OpenAI and Plaintiffs or the Members of the
18 Subclasses regarding use of these individuals’ information, and no effective consent.

19 229. Plaintiffs and the User Subclasses were neither on constructive notice nor inquiry
20 notice of the privacy policy on the ChatGPT platform.

21 ***2. Defendants’ Use of Consumer Data Far Exceeds Industry Standards and***
22 ***their Own Representations***

23 230. The Federal Trade Commission has promulgated numerous guides for businesses
24 highlighting the importance of implementing reasonable data security practices. According to the
25 FTC, the need for data security should be factored into all decision-making.²³⁰

26 231. In 2016, the FTC updated its publication, Protecting Personal Information: A Guide

27 _____
28 ²³⁰ *Start with Security: A Guide for Business: Lessons Learned from FTC Cases*, FED. TRADE
COMM’N. (June, 2015), [https://www.ftc.gov/system/files/documents/plain-language/pdf0205-
startwithsecurity.pdf](https://www.ftc.gov/system/files/documents/plain-language/pdf0205-startwithsecurity.pdf).

1 for Business, which established cybersecurity guidelines for businesses.²³¹ The guidelines note that
2 businesses should protect the personal customer information that they keep; properly dispose of
3 personal information that is no longer needed; encrypt information stored on computer networks;
4 understand their network’s vulnerabilities; and implement policies to correct any security problems.

5 232. The FTC further recommends that entities not maintain personally identifiable
6 information longer than is needed for authorization of a transaction; limit access to sensitive data;
7 require complex passwords to be used on networks; use industry-tested methods for security;
8 monitor for suspicious activity on the network; and verify that third-party service providers have
9 implemented reasonable security measures. The FTC has brought enforcement actions against
10 entities engaged in commerce for failing to adequately and reasonably protect customer data,
11 treating the failure to employ reasonable and appropriate measures to protect against unauthorized
12 access to confidential consumer data as an unfair act or practice prohibited by Section 5 of the
13 Federal Trade Commission Act (“FTCA”), 15 U.S.C. § 45. Orders resulting from these actions
14 further clarify the measures businesses must take to meet their data security obligations.

15 233. Defendants fail to meet these obligations, as they directly feed consumers’ personal
16 information into their LLMs for training purposes.

17 234. Even if the click-through button discussed above could constitute a binding
18 agreement—it cannot—the substance of the policies is insufficient to put any consumer on notice
19 of what to expect with regard to the use of their information. The policies lay out vague promises
20 regarding how and when the users’ data can and will be shared, and affirm that all laws are being
21 complied with—even where such affirmations are internally inconsistent.²³² For example, under the
22 heading “Additional U.S. State Disclosures,” the Privacy Policy lists five different categories of
23 “Personal Information,” including one category that OpenAI identifies as “Sensitive Personal
24 Information,” and states that OpenAI discloses information from *all five* of the various categories
25 to “our affiliates, vendors and service providers, law enforcement, and parties involved in

26 ²³¹ *Protecting Personal Information: A Guide for Business*, FED. TRADE COMM’N. (Oct. 2016),
27 [https://www.ftc.gov/system/files/documents/plain-language/pdf-0136_proteting-personal-](https://www.ftc.gov/system/files/documents/plain-language/pdf-0136_proteting-personal-information.pdf)
information.pdf.

28 ²³² *Privacy Policy*, OPENAI, <https://openai.com/policies/privacy-policy> (last updated June 23,
2023).

1 Transactions.” Yet a few paragraphs down, the policy then inexplicably asserts “We don’t sell
2 Personal Information or share Personal Information.” No explanation is given as to what is meant
3 by the assertion that the company both *does* and *does not* share Personal Information.

4 235. As of June 23, 2023, Defendants changed this language to clarify that they “don’t
5 ‘sell’ Personal Information or ‘share’ Personal Information for cross-contextual behavioral
6 advertising (as those terms are defined under applicable local law).”²³³ Nevertheless, no explanation
7 is given as to how Defendants can ensure that the entities with which they are sharing users’ personal
8 information with are not, in fact, using it for cross-contextual behavior advertising. Defendants also
9 do not disclose the specific purposes for which they do use such sensitive data.

10 236. Moreover, the Policy alerts consumers that to the extent local law entitles them to
11 request deletion of their Personal Information, they can exercise this right (amongst others) by
12 sending a request to dsar@openai.com. Yet nothing in the privacy policy explains that information
13 which has already been incorporated into Defendants’ LLMs *can never really* be removed.

14 237. Finally, even if users are on notice of the Privacy Policy (and they are not), the Privacy
15 Policy does not disclose wiretapping. There is **zero** adequate consent for wiretapping, and
16 OpenAI’s terms and conditions are convoluted, inconspicuous, and consist of numerous documents,
17 impossible to decipher by reasonable consumers. There are no conspicuous or clear disclosures that
18 all conversations are wiretapped, recorded, and shared with numerous entities—none of which are
19 disclosed.

20 238. Beyond Defendants’ legal obligations to protect the confidentiality of individuals’
21 User Data, Defendants’ privacy policy and online representations affirmatively and unequivocally
22 state that any personal information provided to Defendants will remain secure and protected. Since
23 ChatGPT’s inception, Defendants have represented and continue to represent that:

24 “We at OpenAI OpCo, LLC (together with our affiliates, “OpenAI”, “we”,
25 “our” or “us”) respect your privacy and are strongly committed to keeping
secure any information we obtain from you or about you.”

26 “We implement commercially reasonable technical, administrative, and
27 organizational measures to protect Personal Information both online and
offline from loss, misuse, and unauthorized access, disclosure, alteration, or

28 _____
²³³ *Id.*

destruction.”

“OpenAI does not knowingly collect Personal Information from children under the age of 13.”²³⁴

239. Defendants have failed to adhere to a single promise vis-à-vis their duty to safeguard User Data. Defendants have made these privacy policies and commitments available in ChatGPT. In these representations to Plaintiffs and Class Members and the public, Defendants promised to take specific measures to protect its members’ information, consistent with industry standards and federal and state law. However, they did not.

240. Plaintiffs and Class Members relied to their detriment on Defendants’ uniform representations and omissions regarding data security. Now that their sensitive personal and medical information is in the possession of third parties, Plaintiffs and Class Members face a constant threat of continued harm. Collection of such sensitive information without consent or notice poses a great threat to individuals by subjecting them to the danger of potential attacks and embarrassment.

241. Plaintiffs and Class Members trusted Defendants’ Products when inputting sensitive and valuable User Data. Had Defendants disclosed to Plaintiffs and its other members that every click, every search, and every input of sensitive information was being tracked, recorded, collected, and disclosed to third parties—Plaintiffs would not have trusted Defendants’ Products to input such sensitive information.

242. Defendants knew or should have known that Plaintiffs and Class Members would reasonably rely upon, and trust Defendants’ promises regarding security and safety of its data and systems.

243. Additionally, Defendants were aware that ChatGPT collects, tracks, and discloses Plaintiffs’ and Class Members’ User Data, including sensitive information.

244. By virtue of how ChatGPT is “trained,” i.e., through the collection and processing of a massive corpus of data, Defendants were aware that their Users’ data would be collected and disclosed to third parties every time a user interacted with ChatGPT.

CLASS ALLEGATIONS

245. **Class Definition:** Plaintiffs bring this action pursuant to Federal Rules of Civil

²³⁴ *Id.*

1 Procedure Sections 23(b)(2), 23(b)(3), and 23(c)(4), on behalf of Plaintiffs and the Class defined as
2 follows:

- 3 a. **Non-User Class:** All persons in the United States whose PII, Personal
4 Information, or Private Information was disclosed to, or accessed, collected,
5 tracked, taken, or used by Defendants without consent or authorization.
- 6 b. **ChatGPT User Class:** All persons in the United States who used ChatGPT,
7 whose Private Information was disclosed to, or intercepted, accessed, collected,
8 tracked, taken, or used by Defendants without consent or authorization.
- 9 c. **ChatGPT API User Class:** All persons in the United States who used other
10 platforms, programs, or applications which integrated ChatGPT technology,
11 whose Private Information was disclosed to, or intercepted, accessed, collected,
12 tracked, taken, or used by Defendants without consent or authorization.
- 13 d. **Microsoft User Class:** All persons in the United States who used Microsoft
14 platforms, programs, or applications which integrated ChatGPT technology,
15 whose Private Information was disclosed to, or intercepted, accessed, collected,
16 tracked, taken, or used by Defendants without consent or authorization.
- 17 e. **ChatGPT Plus User Class:** All persons in the United States who used Chat-
18 GPT website or mobile app and whose Personal Information or PII was
19 intercepted, accessed, collected, tracked, stored, shared, taken, or used by
20 Defendants without consent and/or authorization.

21 **State-Wide Subclasses:**

22 **The California Subclasses**

- 23 i. **California Non-User SubClass:** All persons within the State of
24 California whose PII, Personal Information, or Private Information
25 was disclosed to, or accessed, collected, tracked, taken, or used by
26 Defendants without consent or authorization.
- 27 ii. **California ChatGPT User SubClass:** All persons within the State
28 of California who used ChatGPT, whose Private Information was
disclosed to, or intercepted, accessed, collected, tracked, taken, or
used by Defendants without consent or authorization.
- iii. **California ChatGPT Plus User SubClass:** All persons within the
State of California who used Chat-GPT website or mobile app and
whose Personal Information or PII was intercepted, accessed,
collected, tracked, stored, shared, taken, or used by Defendants
without consent and/or authorization.

The New York Subclasses

- i. **New York Non-User SubClass:** All persons within the State of
New York whose PII, Personal Information, or Private Information
was disclosed to, or accessed, collected, tracked, taken, or used by
Defendants without consent or authorization.
- ii. **New York ChatGPT User SubClass:** All persons within the State

1 of New York who used ChatGPT, whose Private Information was
2 disclosed to, or intercepted, accessed, collected, tracked, taken, or
3 used by Defendants without consent or authorization.

4 iii. **New York ChatGPT Plus User SubClass:** All persons within the
5 State of New York who used Chat-GPT website or mobile app and
6 whose Personal Information or PII was intercepted, accessed,
7 collected, tracked, stored, shared, taken, or used by Defendants
8 without consent and/or authorization.

9 246. **The following people are excluded from the Classes and Subclasses:** (1) any Judge
10 or Magistrate presiding over this action and members of their judicial staff and immediate families;
11 (2) Defendants, Defendants' subsidiaries, parents, successors, predecessors, and any entity in which
12 the Defendants or their parents have a controlling interest and its current or former officers and
13 directors; (3) persons who properly execute and file a timely request for exclusion from the Class;
14 (4) persons whose claims in this matter have been finally adjudicated on the merits or otherwise
15 released; (5) Plaintiffs' counsel and Defendants' counsel; and (6) the legal representatives,
16 successors, and assigns of any such excluded persons.

17 247. Plaintiffs reserve the right under Federal Rule of Civil Procedure 23 to amend or
18 modify the Class to include a broader scope, greater specificity, further division into subclasses, or
19 limitations to particular issues. Plaintiffs reserve the right under Federal Rule of Civil Procedure
20 23(c)(4) to seek certification of particular issues.

21 248. The requirements of Federal Rules of Civil Procedure 23(a), 23(b)(2), and 23(b)(3)
22 are met in this case.

23 249. The Fed. R. Civ. P. 23(a) elements of Numerosity, Commonality, Typicality, and
24 Adequacy are all satisfied.

25 250. **Ascertainability:** Membership of the Classes and Subclasses is defined based on
26 objective criteria and individual members will be identifiable from Defendants' records, records of
27 third-party platforms/applications which integrate ChatGPT, including the massive data storage,
28 consumer accounts, and enterprise services that Defendants offer. Identification is also available
through self-identification methods.

251. **Numerosity:** The precise number of the Members of Classes and Subclasses is not
available to Plaintiffs, but individual joinder is demonstrably impracticable.

1 252. **Commonality:** Commonality requires that the Members of Classes and Subclasses
 2 allege claims which share common contention such that determination of its truth or falsity will
 3 resolve an issue that is central to the validity of each claim in one stroke. Here, there is a common
 4 contention for all Classes and Subclasses are as follows:

5 **Defendants' Web-Scraping Practices (Non-User Class)**

- 6 a) Whether the members of Non-User Class had a protected property right in their data;
- 7 b) Whether Defendants scraped the protected data belonging to Non-User Class
 8 members without consent;
- 9 c) Whether Defendants' collection, scraping, and uses of the protected Non-User Class
 10 Members of protected data violates:
- 11 3. Electronic Communication Privacy Act, 18 U.S.C. §§ 2510, *et. seq.*
- 12 4. Computer Fraud and Abuse Act, 18 U.S.C. §§ 1030, *et. seq.*
- 13 5. California Constitution right to privacy;
- 14 6. California Invasion of Privacy Act, Cal. Pen. Code §§ 630, *et seq.*
- 15 7. California Unfair Competition Law, Bus. & Prof Code § 17200;
- 16 8. New York General Business Law §§ 349, *et seq.*
- 17 d) Whether Defendants' collection, scraping, and uses of the protected Non-User Class
 18 Members of protected data constitutes:
- 19 1. Common law Negligence;
- 20 2. Unlawful Intrusion upon Seclusion under California laws;
- 21 3. Conversion;
- 22 4. Larceny/Receipt of Stolen Property under Cal. Pen. Code § 496(a) and (c).
- 23 e) Whether as a result of Defendants' collection, scraping, and uses of the protected
 24 Non-User Class Members of protected data, Non-User Class Members suffered
 25 monetary damages, including but not limited to actual damages, statutory damages,
 26 punitive damages, treble damages, or other monetary damages.
- 27 f) Whether as a result of Defendants' collection, scraping, and uses of the protected
 28 Non-User Class Members of protected data, Non-User Class Members are entitled

1 to equitable relief, including but not limited to restitution, disgorgement of profits,
2 injunctive and declaratory relief, or other equitable remedies.

3 **Defendants’ Collection/Interception Practices of Private Information From ChatGPT**
4 **User, ChatGPT Plug-In User, ChatGPT Plus User Classes, and Subclasses:**

- 5 a) Whether Defendants failed to advise the members of Classes and Subclasses the
6 extent to which Defendants intercepted, received, collected Private Information;
- 7 b) Whether Defendants intercepted, received, or collected communications, tracked all
8 activities, chat history, and other Private Information from the Users of Other
9 Platforms Which Integrate ChatGPT without consent of such Users.
- 10 c) Whether Microsoft Defendant intercepted, received, or collected communications,
11 tracked all activities, chat history, and other Private Information of ChatGPT Users,
12 without consent of such Users;
- 13 d) Whether Open AI Defendant aided, abetted, and otherwise conspired with Microsoft
14 Defendant, to allow Defendant Microsoft’s interception, receipt, or collection of
15 communications, tracking of all activities, and other Private Information of
16 ChatGPT Users, without consent of such Users;
- 17 e) Whether Defendants’ conduct of intercepting, receipt, collection of Private
18 Information of the members of Classes and Subclasses violated federal and state
19 privacy laws, anti-wiretapping laws, or other tort laws, including but not limited to:
- 20 1. Electronic Communication Privacy Act, 18 U.S.C. § 2510 *et. seq.*
21 2. Computer Fraud and Abuse Act, 18 U.S.C. § 1030 *et. seq.*
22 3. California Constitution right to privacy;
23 4. California Invasion of Privacy Act, Cal. Pen. Code §§ 630 *et seq.*
24 5. California Unfair Competition Law, Bus. & Prof Code §§ 17200;
25 6. Common law Negligence;
26 7. Unlawful Intrusion upon Seclusion under California laws;
27 8. Conversion.
- 28 f) Whether as a result of Defendants’ collection, scraping, and uses of the protected

1 Private Information, ChatGPT User, ChatGPT Plug-In User, ChatGPT Plus User
2 Class Members and Subclass Members suffered monetary damages, including but
3 not limited to actual damages, statutory damages, punitive damages, treble damages,
4 or other monetary damages.

5 g) Whether as result of Defendants' interception, collection, receipt, or unauthorized
6 uses of Private Information, ChatGPT User, ChatGPT Plug-In User, ChatGPT Plus
7 User Class Members and Subclass Members are entitled to equitable relief,
8 including but not limited to restitution, disgorgement of profits, injunctive and
9 declaratory relief, or other equitable remedies.

10 253. **Typicality:** Plaintiffs' claims are typical of the claims of other Class Members in that
11 Plaintiffs and the Class Members sustained damages arising out of Defendants' uniform wrongful
12 conduct and data collecting practices, interception/sharing of the collected data with each other, and
13 use of such data in attempt to train the AI Products, and further develop the Products.

14 254. **Adequate Representation:** Plaintiffs will fairly and adequately represent and protect
15 the interests of the Members of Classes and Subclasses. Plaintiffs' claims are made in a
16 representative capacity on behalf of the Members of Classes and Subclasses. Plaintiffs have no
17 interests antagonistic to the interests of the other Members of Classes and Subclasses. Plaintiffs
18 have retained competent counsel to prosecute the case on behalf of Plaintiffs and the Class. Plaintiffs
19 and Plaintiffs' counsel are committed to vigorously prosecuting this action on behalf of the
20 Members of Classes and Subclasses.

21 255. **This case also satisfies Fed. R. Civ. P. 23(b)(3) - Predominance:** There are many
22 questions of law and fact common to the claims of Plaintiffs and Members of Classes and
23 Subclasses, and those questions predominate over any questions that may affect individual Class
24 Members. Common questions and/or issues for Class members include the questions listed above
25 in *Commonality*, and also include, but are not necessarily limited to the following:

- 26 a) Whether Defendants violated the California Invasion of Privacy Act;
27 b) Whether Defendants' unauthorized disclosure of Users' sensitive information was
28 negligent;

- 1 c) Whether Defendants owed a duty to Plaintiffs' and Class Members not to disclose
- 2 their sensitive user information to unauthorized third parties;
- 3 d) Whether Defendants breached their duty to Plaintiffs' and Class Members not to
- 4 disclose their sensitive user information to unauthorized third parties;
- 5 e) Whether Defendants represented to Plaintiffs and the Class that they would protect
- 6 Plaintiffs' and the Members of Classes and Subclasses Private Information;
- 7 f) Whether Defendants violated Plaintiffs' and Class Members' right to privacy;
- 8 g) Whether Plaintiffs and Class members are entitled to actual damages, enhanced
- 9 damages, statutory damages, restitution, disgorgement, and other monetary
- 10 remedies provided by equity and law;
- 11 h) Whether Defendants' conduct was unlawful or deceptive;
- 12 i) Whether Defendants were unjustly enriched by their conduct under the laws of
- 13 California.
- 14 j) Whether Defendants fraudulently concealed their conduct; and
- 15 k) Whether injunctive and declaratory relief and other equitable relief is warranted.

16 256. **Superiority:** This case is also appropriate for class certification because class
17 proceedings are superior to all other available methods for the fair and efficient adjudication of this
18 controversy as joinder of all parties is impracticable. The damages suffered by individual Members
19 of Classes and Subclasses will likely be relatively small, especially given the burden and expense
20 of individual prosecution of the complex litigation necessitated by Defendants' actions. Thus, it
21 would be virtually impossible for the individual Members of Classes and Subclasses to obtain
22 effective relief from Defendants' misconduct. Even if Class Members could mount such individual
23 litigation, it would still not be preferable to a class action, because individual litigation would
24 increase the delay and expense to all parties due to the complex legal and factual controversies
25 presented in this Complaint. By contrast, a class action presents far fewer management difficulties
26 and provides the benefits of single adjudication, economy of scale, and comprehensive supervision
27 by a single Court. Economies of time, effort, and expense will be enhanced, and uniformity of
28 decisions ensured.

1 257. Likewise, particular issues under Rule 23(c)(4) are appropriate for certification
2 because such claims present only particular, common issues, the resolution of which would advance
3 the disposition of this matter and the parties' interests therein.

4 **CALIFORNIA LAW SHOULD APPLY TO OUT-OF-STATE PLAINTIFF'S & CLASS**
5 **MEMBERS' NON-STATUTORY CLAIMS**

6 258. Courts "have permitted the application of California law where the plaintiffs' claims
7 were based on alleged misrepresentations [or misconduct] that were disseminated from
8 California." *Ehret v. Uber Technologies, Inc.*, 68 F. Supp. 3d 1121, 1130 (N.D. Cal.
9 2014). "California courts have concluded that state statutory remedies may be invoked by out-of-
10 state parties when they are harmed by wrongful conduct occurring in California." *In re iPhone 4S*
11 *Consumer Litig.*, No. C 12-1127 CW, 2013 WL 3829653, at *7 (N.D. Cal. July 23, 2013) (internal
12 quotation marks and citation omitted).

13 259. This is particularly true for non-statutory claims where the defendant has a choice-of-
14 law provision that applies California law to that defendant's conduct.

15 260. However, there is sound public policy to allow statutory claims from other states to
16 proceed against a defendant regardless of that defendant's choice of law provision. *See, e.g., In re*
17 *Facebook Biometric Info. Priv. Litig.*, 185 F. Supp. 3d 1155, 1168–70 (N.D. Cal. 2016).

18 261. Defendant Open AI is headquartered in California; this is where Defendant Open AI's
19 nerve center of its business operations is located. This is where Defendant Open AI has its high-
20 level officers direct, control, coordinate, and manage its activities, including policies, practices,
21 research and development, and other decisions affecting Defendants' Products. This is where the
22 majority of unlawful conduct took place – from development of the AI products, decisions
23 concerning AI Products and training of the AI, web scraping practices, and other major decisions
24 which affected all Class Members. Furthermore, Defendant Microsoft operates in the state of
25 California. Upon information and belief, decisions concerning Defendants' Products were entered
26 into in California.

27 262. Furthermore, Defendant Open AI requires that California law applies to disputes
28 between Defendant Open AI and ChatGPT Users.

 263. The State of California, therefore, has significant interests to protect all residents and

1 citizens of the United States against a company headquartered and doing business in California, and
2 has a greater interest in the claims of Plaintiff J.H. and the Classes than any other state, and the state
3 most intimately concerned with the claims and outcome of this litigation.

4 264. California has significant interest in regulating the conduct of businesses operating
5 within its borders, and that California has the most significant relationship with Defendants – as
6 Defendant Open AI is headquartered in California, and Defendant Microsoft conducts business (at
7 least as it relates to Defendant Open AI) in California, there is no conflict in applying California
8 law to non-resident consumer claims.

9 265. Excluding out-of-state statutory claims, application of California law to the Classes’
10 claims is neither arbitrary nor fundamentally unfair because choice of law principles applicable to
11 this action support the application of California law to the nationwide claims of all Class Members.

12 266. Application of California law to Defendants is consistent with constitutional due
13 process.

14 **COUNT ONE: VIOLATION OF ELECTRONIC COMMUNICATIONS PRIVACY ACT,**
15 **18 U.S.C. § 2510, et seq.**
16 **(on behalf of ChatGPT, ChatGPT API User, Microsoft User Classes against Defendants)**

17 267. Plaintiffs hereby incorporate Paragraphs 1 through 266 as if fully stated herein.

18 268. The Federal Wiretap Act, as amended by the Electronic Communications Privacy Act
19 of 1986 (the “Wiretap Act”), prohibits the intentional interception of the contents of any wire, oral,
20 or electronic communication through the use of a device. 18 U.S.C. § 2511.

21 269. The following constitute “devices” within the meaning of the Wiretap Act, 18 U.S.C.
§ 2510(5):

- 22 a. The computer codes and programs that Defendants use to track the Plaintiffs’
23 and Class members’ communications;
- 24 b. The Plaintiffs’ and Class members’ browsers and applications;
- 25 c. The Plaintiffs’ and Class members’ computing and mobile devices;
- 26 d. Defendants’ web servers;
- 27 e. The web servers of websites from which Defendants tracked and intercepted
28 the Plaintiffs’ and Class members’ communications;

1 f. The computer codes and programs used by Defendants to effectuate their
2 tracking and interception of the Plaintiffs' and Class members'
3 communications;

4 g. The plan that Defendants carried out to effectuate its tracking and interception
5 of the Plaintiffs' and Class members' communications.

6 270. The Wiretap Act protects both the sending and reception of communications.

7 271. The Wiretap Act provides a private right of action to any person whose wire, oral, or
8 electronic communication is intercepted. 18 U.S.C. § 2520(a).

9 272. Defendants' actions in tracking and intercepting users' communications were
10 intentional. On information and belief, Defendants are aware that they are tracking and intercepting
11 these communications as outlined in this complaint and they have taken no remedial actions.

12 273. Defendants' actions were done contemporaneously with the Plaintiffs' and Class
13 members' sending and receiving those communications.

14 274. Defendants' interception included "contents" of electronic communications made
15 from Plaintiffs and Class members to websites and other web properties other than Defendants' in
16 the form of detailed URL requests, webpage browsing histories, search queries, and other
17 information that Plaintiffs and Class members sent to those websites and for which Plaintiffs
18 received communications in return from those websites.

19 275. The transmission of data between Plaintiffs and Class members on the one hand and
20 the websites and other web properties other than Defendants' on which Defendants tracked and
21 intercepted Plaintiffs' and Class members' communications on the other, without authorization were
22 "transfer[s] of signs, signals, writing, . . . data, [and] intelligence of [some] nature transmitted in
23 whole or in part by a wire, radio, electromagnetic, photoelectronic, or photooptical system that
24 affects interstate commerce[.]" and therefore qualify as "electronic communications" within the
25 meaning of the Wiretap Act. 18 U.S.C. § 2510(12).

26 276. Defendants, in their conduct alleged herein, were not providing an "electronic
27 communication service," as that term is defined in 18 U.S.C. § 2510(12) and is used elsewhere in
28 the Wiretap Act. Defendants were not acting as an Internet Service Provider and the conduct alleged

1 herein does not arise from their provision of separate lines of business.

2 277. None of the Defendants were authorized parties to the communications because
3 Plaintiffs and Class members were unaware of the collection and interception. Neither can
4 Defendants manufacture their own status as parties to the communications by surreptitiously
5 intercepting those communications.

6 278.

7
8 279. Both Defendants had a tortious and/or criminal intent in (a) obtaining the Private
9 Information, (b) sharing the Private Information with each other; (c) feeding the Private Information
10 into the Products, to train, develop, and commercialize their Products. Their actions were knowing
11 and deliberate, especially since Defendants were well aware that consumers did not want nor allow
12 Defendants to use their Private Information for training of the Products.

13 280. **Electronic Communications.** Electronic communication means any “transfer[s] of
14 signs, signals, writing, . . . data, [and] intelligence of [some] nature transmitted in whole or in part
15 by a wire, radio, electromagnetic, photoelectronic, or photooptical system that affects interstate
16 commerce.” 18 U.S.C. § 2510(12). Here, the following communications qualify as
17 “communications” under the ECPA:

- 18 a) **Communications On ChatGPT:** Plaintiffs’ and Class Members’ communications
19 (including but not limited to chats, comments, replies, searches, keystrokes, signals,
20 mouse clicks, or other data, activity, or intelligence) on ChatGPT intercepted by
21 Defendant Microsoft;
- 22 b) **ChatGPT Intercepted Communications On Platforms Which Integrated**
23 **ChatGPT API:** Plaintiffs’ and Class Members’ communications (including but not
24 limited to chats, comments, replies, searches, keystrokes, signals, mouse clicks, or
25 other data, activity, or intelligence) on various applications, platforms, or websites
26 which integrate ChatGPT API (i.e. Stripe, Snapchat, etc.) intercepted by
27 Defendants;
- 28 c) **Communications on Microsoft Platforms:** Plaintiffs’ and Class Members’
communications (including but not limited to chats, comments, replies, searches,
keystrokes, mouse clicks, signals, or other data, activity, or intelligence) on
Microsoft platforms which integrate ChatGPT API (i.e. Microsoft Teams, Outlook,
etc.) intercepted by Defendant Open AI;

1 281. **Content.** The ECPA defines content, when used with respect to electronic
2 communications, to “include [] any information concerning the substance, purport, or meaning of
3 that communication.” 18 U.S.C. § 2510(8).

4 282. Plaintiffs, and the members of all Classes and Subclasses have an expectation of
5 privacy in their communications, entered keystrokes, chats, comments, replies, searches, signals,
6 and other data, activity, or intelligence, and they exercised a reasonable expectation of privacy
7 concerning the transmission of that content.

8 283. **Interception.** The ECPA defines interception as the “acquisition of the contents of
9 any wire, electronic, or oral communication through the use of any electronic, mechanical, or other
10 device” and “contents . . . include [] any information concerning the substance, purport, or meaning
11 of that communication.” 18 U.S.C. §§ 2510(4), (8).

12 284. Defendants intentionally accessed, and obtained access to the contents of Plaintiffs’,
13 the Classes’, and Subclasses’ protected computers and obtained information concerning the
14 substance, purport, or meaning of communications, thereby, and in doing so, exceeded authority
15 granted by Plaintiffs, the Classes, and Subclasses to access the protected computers.

16 285. **Electronic Communication Service.** The ECPA defines electronic communication
17 service as “any service which provides to users thereof the ability to send or receive wire or
18 electronic communications.” 18 U.S.C § 2510(15). The following services constitute “electronic
19 communication services:”

- 20 (1) Reddit, Twitter, YouTube, Spotify, TikTok, and other websites which were scraped
21 by Defendants;
- 22 (2) Third Party websites, programs, and applications, which integrate ChatGPT
23 technology;
- 24 (3) Microsoft platforms, programs, applications, and websites, which integrate
25 ChatGPT technology;
- 26 (4) Open AI website and mobile application(s) for ChatGPT.

27 286. **Electronic, Mechanical, or Other Device.** The ECPA defines “electronic,
28 mechanical, or other device” as “any device...which can be used to intercept a[n]...electronic

1 communication[.]” 18 U.S.C. § 2510(5). The following constitute “devices” within the meaning of
2 18 U.S.C. § 2510(5):

- 3 (1) Plaintiffs’ and Classes’, Subclasses’ computing devices (Mac and Windows devices
4 present on computers, mobile phones, tablets, or other devices);
- 5 (2) Plaintiffs’ and Classes’, Subclasses’ browsers;
- 6 (3) Defendants’ web-servers, platforms, and applications;
- 7 (4) Third-Party web-servers, platforms, and applications, where ChatGPT API
8 technology was implemented;
- 9 (5) The tracking codes deployed by Defendants to effectuate the sending and acquisition
10 of communications.

11 **I. Interception of Communications Between ChatGPT API Class Members which**
12 **occurred on Third-Party Websites, Platforms, Applications, Programs which have**
13 **integrated ChatGPT API. [Microsoft User Class is Excluded]**

14 287. The allegations for violation of 18 U.S.C. § 2510 arising out of Defendants’
15 interception of Plaintiffs’, and ChatGPT API Class Members’ (collectively referred to as ChatGPT
16 API Class Members) communications which occurred on various applications, platforms, websites
17 which integrate ChatGPT technology (i.e., Stripe, Snapchat, etc.).

18 288. The transmissions of Plaintiffs’, and ChatGPT API Class Members’ communications
19 (including but not limited to chats, comments, replies, searches, keystrokes, mouse
20 clicks/movements, signals, browser activity, or other data, activity, or intelligence) on various
21 applications, programs, platforms, and websites which integrate ChatGPT technology (i.e., Stripe,
22 Snapchat, etc.) qualify as “communications” under 18 U.S.C. § 2510(12).

23 289. By integrating ChatGPT technology on third party platforms, Defendants are in the
24 unique position of having unrestricted, real-time access to the users’ every input, move, mouse click,
25 chat, comment, reply, search, keystroke, browser activity, or other data, activity, or intelligence on
26 the third-party platform.

27 290. As Plaintiffs and ChatGPT API Class Members interact with each other or the third-
28 party entities, Defendants intentionally tap, electrically or otherwise intercept, the lines of internet

1 communications between Plaintiffs and ChatGPT API Class Members, and/or third-party entities.

2 291. In disregard for Plaintiffs', and ChatGPT API Class Members' privacy rights,
3 Defendants act as a third-party "eavesdropper," redirecting Plaintiffs and ChatGPT API Class
4 Members' electronic communications to Defendants' own servers for appropriation, and training of
5 their Products.

6 292. Defendants' interception of the contents of Plaintiffs' and ChatGPT API Class
7 Members' communications happens contemporaneously with their exchange of such
8 communications, whether such communications are directed to Plaintiffs' and ChatGPT API Class
9 Members' friends, colleagues, or third-party entities. As described above, the ChatGPT API is
10 designed to simultaneously intercept and send a recording of each keystroke, mouse click,
11 movement, writing, or other data, activity, or intelligence to Defendants sufficient to not only
12 identify Plaintiffs and ChatGPT API Class Members also to be able to understand, collect, and use
13 for training Plaintiffs' and ChatGPT API Class Members' communications.

14 293. **Unauthorized Purpose.** Plaintiffs and ChatGPT API Class Members did not
15 authorize Defendants to acquire, access, or intercept the content of their communications on third
16 party platforms, websites, applications. Therefore, such interception and recording of
17 communications invades Plaintiffs', and ChatGPT API Class Members' privacy. Defendants
18 intentionally intercepted the contents of Plaintiffs' and ChatGPT API Class Members' electronic
19 communications for the purpose of committing a tortious act in violation of the Constitution or laws
20 of the United States or of any State – namely, the knowing intrusion into a private place,
21 conversation, or matter that would be highly offensive to a reasonable person.

22 294. **While in Transmission.** Through this calculated scheme of using ChatGPT API to
23 intercept, acquire, transmit, and record Plaintiffs' and ChatGPT API Class Members' electronic
24 communications, Defendants willfully and without valid consent from all parties to the
25 communication, take unauthorized measures to read and understand the contents or meaning of the
26 electronic communications of Plaintiffs, and ChatGPT API Class. The interception and recording
27 of electronic communications occurs while the electronic communications are in transit or passing
28 over any wire, line, or cable, or are being sent from or received at any place.

1 295. In sending and in acquiring the content of Plaintiffs’, and ChatGPT API Class
2 Members’ communications with third-party platforms, Defendants’ purpose was tortious, and
3 designed to violate federal and state legal laws. By intentionally using, or endeavoring to use, the
4 contents of the electronic communications of Plaintiffs, ChatGPT API Class and Subclass Members,
5 while knowing or having reason to know that the information was obtained through the interception
6 of an electronic communication, Defendants violate 18 U.S.C. § 2511(1)(a).

7 296. Plaintiffs, individually, on behalf of the GPT API Class and Subclass Members, seek
8 all monetary and non-monetary relief allowed by law, including actual damages, statutory damages,
9 punitive damages, preliminary and other equitable or declaratory relief, and attorneys’ fees and
10 costs.

11 **II. Microsoft’s Interception of Communications Between ChatGPT Class Members**

12 297. The allegations for violation of 18 U.S.C. § 2510 arising out of Defendant Microsoft’s
13 interception of Plaintiffs, ChatGPT User Class Members’ communications which occurred on
14 ChatGPT platform.

15 298. The transmissions of Plaintiffs’, ChatGPT User Class Members’ communications
16 (including but not limited to chats, comments, replies, searches, keystrokes, mouse
17 clicks/movements, signals, browser activity, or other data, activity, or intelligence) on ChatGPT
18 platform qualify as “communications” under 18 U.S.C. § 2510(12).

19 299. By integrating ChatGPT technology on third party platforms, Defendants are in the
20 unique position of having unrestricted, real-time access to the users’ every input, move, mouse click,
21 chat, comment, reply, search, keystroke, browser activity, or other data, activity, or intelligence on
22 the third-party platform.

23 300. As Plaintiffs, ChatGPT User Class Members’ interact with each other or the third-
24 party entities, Defendant Open AI intentionally divulges and Defendant Microsoft intentionally
25 taps, electrically or otherwise intercepts the lines of internet communications between Plaintiffs,
26 ChatGPT, and/or third party entities (integrated within ChatGPT through plug-in technologies).

27 301. In disregard for Plaintiffs’ and ChatGPT User Class Members’ privacy rights,
28 Defendant Microsoft acts as a third-party “eavesdropper,” redirecting Plaintiffs’ and ChatGPT User

1 Class Members' electronic communications to Defendant Microsoft's own servers for
2 appropriation, and training of their Products.

3 302. Defendant Microsoft's interception of the contents of Plaintiffs', ChatGPT User Class
4 Members' communications happens contemporaneously with their exchange of such
5 communications, whether such communications are directed to Defendant Open AI or third-party
6 entities. As described above, the ChatGPT is designed to simultaneously intercept and send a
7 recording of each keystroke, mouse click, movement, writing, or other data, activity, or intelligence
8 to Defendant Microsoft sufficient to not only identify Plaintiffs, and ChatGPT User Class Members,
9 but also to be able to understand, collect, and use for training Plaintiffs' and ChatGPT User Class
10 Members' communications.

11 303. **Unauthorized Purpose.** Plaintiffs and ChatGPT User Class Members did not
12 authorize Defendant Microsoft to acquire, access, or intercept the content of their communications
13 on third party platforms, websites, applications. Moreover, Plaintiffs and ChatGPT User Class
14 Members did not authorize either Defendant to train their AI Products on private information
15 acquired by Defendants. Therefore, such interception and recording of communications invades
16 Plaintiffs', ChatGPT User Class Members' privacy. Defendant Open AI illegally divulged the
17 content of such communications to Defendant Microsoft. Defendant Microsoft intentionally
18 intercepted the contents of Plaintiffs' and ChatGPT User Class Members' communications for the
19 purpose of committing a tortious act in violation of the Constitution or laws of the United States or
20 of any State – namely, the knowing intrusion into a private place, conversation, or matter that would
21 be highly offensive to a reasonable person.

22 304. **While in Transmission.** Through this calculated scheme of using ChatGPT
23 technology to intercept, acquire, transmit, and record Plaintiffs', and ChatGPT User Class
24 Members' electronic communications, Defendant Microsoft willfully and without any iota of valid
25 consent from all parties to the communication, takes unauthorized measures to read and understand
26 the contents or meaning of the electronic communications of Plaintiffs and ChatGPT User Class
27 Members. The interception and recording of electronic communications occur while the electronic
28 communications are in transit or passing over any wire, line, or cable, or are being sent from or

1 received at any place.

2 305. In sending and in acquiring the content of Plaintiffs’, and Class Members’
3 communications with third-party platforms, Defendants’ purpose was tortious, and designed to
4 violate federal and state laws. By intentionally using, or endeavoring to use, the contents of the
5 electronic communications of Plaintiffs, ChatGPT User Class Members, while knowing or having
6 reason to know that the information was obtained through the interception of an electronic
7 communication, Defendant Microsoft violates 18 U.S.C. § 2511(1)(a).

8 306. Plaintiffs, individually, on behalf of the ChatGPT User Class Members, seek all
9 monetary and non-monetary relief allowed by law, including actual damages, statutory damages,
10 punitive damages, preliminary and other equitable or declaratory relief, and attorneys’ fees and
11 costs.

12 **III. Defendant Open AI’s Interception of Microsoft User Class Members which**
13 **occurred on Microsoft’s Websites, Platforms, Applications, Programs which have**
14 **integrated ChatGPT.**

15 307. The allegations for violation of 18 U.S.C. § 2510 arising out of Defendant Open AI’s
16 interception of Microsoft User Class Members’ (collectively “Microsoft Subclasses”)
17 communications with their friends, family, colleagues, or other individuals or third-party entities,
18 which occurred on Microsoft platforms (Teams, Bing, Outlook etc.), which integrate ChatGPT API.

19 308. The transmissions of Plaintiffs’ and Microsoft Subclasses’ communications
20 (including but not limited to chats, comments, replies, searches, keystrokes, signals, mouse
21 clicks/movements, signals, browser activity, or other data, activity, or intelligence) on Microsoft’s
22 various applications, programs, platforms, websites which integrate ChatGPT API qualify as
23 “communications” under 18 U.S.C. § 2510(12).

24 309. By integrating ChatGPT technology within the entire Microsoft suite, Defendant
25 OpenAI is in the unique position of having unrestricted, real-time access to the users’ every input,
26 move, mouse click, chat, comment, reply, search, keystroke, browser activity, or other data, activity,
27 or intelligence.

28 310. As Plaintiffs, Microsoft Subclasses interact with each other or the third-party entities,

1 Defendants intentionally tap, electrically or otherwise intercept, the lines of internet
2 communications between Plaintiffs, Microsoft Subclasses, and/or third-party entities.

3 311. In disregard for Plaintiffs', Microsoft Subclasses Members' privacy rights, Defendant
4 OpenAI acts as a third-party "eavesdropper," redirecting Plaintiffs, Microsoft Subclasses Members'
5 electronic communications to Defendants' own servers for appropriation, and training of their
6 Products.

7 312. Defendant Open AI interception of the contents of Plaintiffs', Microsoft Subclasses
8 Members' communications happens contemporaneously with their exchange of such
9 communications, whether such communications are directed to Plaintiffs', Microsoft Subclasses
10 Members' friends, colleagues, or third-party entities. As described above, the ChatGPT API is
11 designed to simultaneously intercept and send a recording of each keystroke, mouse click, signal,
12 movement, writing, or other data, activity, or intelligence to Defendants sufficient to not only
13 identify Plaintiffs, Microsoft Subclasses Members, but also to be able to understand, collect, and
14 use for training Plaintiffs', Microsoft Subclasses Members' communications.

15 313. **Unauthorized Purpose.** Plaintiffs and Microsoft Subclasses did not authorize
16 Defendant Open AI to acquire, access, or intercept the content of their communications which
17 occurred on Microsoft platforms, applications, programs, and websites. Therefore, such interception
18 and recording of communications invades Plaintiffs', Microsoft Subclasses Members' privacy.
19 Defendant Open AI intentionally intercepted (and continues to intercept) the contents of Plaintiffs',
20 Microsoft Subclasses Members' electronic communications for the purpose of committing a tortious
21 act in violation of the Constitution or laws of the United States or of any State – namely, the knowing
22 intrusion into a private place, conversation, or matter that would be highly offensive to a reasonable
23 person.

24 314. **While in Transmission.** Through this calculated scheme of using ChatGPT API to
25 intercept, acquire, transmit, and record Plaintiffs', Microsoft Subclasses Members' electronic
26 communications, Defendant Open AI willfully and without any iota of valid consent from all parties
27 to the communication, implements unauthorized measures to read and understand the contents or
28 meaning of Plaintiffs' and Microsoft Subclasses' communications. The interception and recording

1 of electronic communications occur while the electronic communications are in transit or passing
2 over any wire, line, or cable, or are being sent from or received at any place.

3 315. In sending and in acquiring the content of Plaintiffs', and Class Members'
4 communications with third-party platforms, Defendant Open AI's purpose was tortious, and
5 designed to violate federal and state laws. By intentionally using, or endeavoring to use, the contents
6 of Plaintiffs' and Microsoft Subclasses' electronic communications, while knowing or having
7 reason to know that the information was obtained through the interception of an electronic
8 communication, Defendant Open AI violated and continues to violate 18 U.S.C. § 2511(1)(a).

9 316. Plaintiffs, individually, on behalf of the Microsoft Subclasses Members, seek all
10 monetary and non-monetary relief allowed by law, including actual damages, statutory damages,
11 punitive damages, preliminary and other equitable or declaratory relief, and attorneys' fees and
12 costs.

13 **COUNT TWO: VIOLATION OF THE COMPUTER FRAUD AND ABUSE ACT, 18 U.S.C.**

14 **§ 1030**

14 **(on behalf of All Plaintiffs against Defendants)**

15 317. Plaintiffs hereby incorporate Paragraphs 1 through 266 as if fully stated herein

16 318. Plaintiffs', the Classes', and Subclasses' computer devices (including but not limited
17 to Mac and Windows devices) were, used for interstate communication and commerce and are
18 therefore "protected computers" under 18 U.S.C. § 1030(e)(2)(B).

19 319. Defendants intentionally accessed Plaintiffs', the Classes and Subclasses Members'
20 protected computers and obtained information thereby, and in doing so exceeded authority granted
21 by Plaintiffs, the Classes, and Subclasses to access the protected computers in violation of 18 U.S.C.
22 § 1030(a)(2)(C). Plaintiffs, the Classes, and Subclasses Members have a civil cause of action for
23 violation of the CFAA under 18 U.S.C. § 1030(g) and have suffered damage or loss.

24 320. **Chat GPT Plug-In:** Defendants owned and operated their Products and ChatGPT
25 Plug-Ins. Defendants integrated ChatGPT Plug-Ins within various platforms, websites, applications,
26 and programs, and thereby intercepted and obtained Plaintiffs', the Classes', and Subclasses' Private
27 Information, inclusive of keywords, mouse clicks, searches, movements, signals, and other activity
28 and intelligence.

1 **321. Microsoft GPT Plug-In:** Defendant Microsoft owned and operated its Microsoft
2 platforms, websites, programs, and applications which integrated Defendants’ ChatGPT Plug-In.
3 Defendant Open AI intercepted and obtained Plaintiffs’, the Classes’, and Subclasses’ Private
4 Information, inclusive of keywords, mouse clicks, searches, movements, signals, and other activity
5 and intelligence. Defendants collected, and transmitted this data to their Products, and used it to
6 train their Products. Defendants’ collected data allows Defendant to determine individual users’
7 precise locations, unique identifiers, cookies, patterns (including browsing patterns, conversational
8 patterns), conversational and browsing activities and habits, and a plethora of other Private
9 Information.

10 **322. ChatGPT:** Defendant Open AI owned and operated its ChatGPT platforms.
11 Defendant Open AI transmits all data from its ChatGPT platforms to Defendant Microsoft;
12 Defendant Microsoft thereby intercepted and obtained Plaintiffs’, the Classes’, and Subclasses’
13 Private Information, inclusive of keywords, mouse clicks, searches, movements, signals, and other
14 activity and intelligence. Defendants collected, and transmitted this data to their Products, and used
15 it to train their Products. Defendants’ collected data allows Defendant to determine individual users’
16 precise locations, unique identifiers, cookies, patterns (including browsing patterns, conversational
17 patterns), conversational and browsing activities and habits, and a plethora of other Private
18 Information.

19 323. Defendants accessed, and otherwise transmitted this data without authorized consent
20 from Plaintiffs, Classes, and Subclasses; or at a minimum, as discussed above, exceed any consent
21 that was given.

22 324. Defendants were actively involved in implementing the unlawful interception alleged
23 herein and promoted the use of their Products to U.S. residents and other companies, knowing about
24 the privacy violations alleged herein. Defendants are also liable for this conduct because it occurred
25 pursuant to the common enterprise of which they are a part.

26 325. Defendants’ conduct caused “loss to 1 or more persons during any 1-year period . . .
27 aggregating at least \$5,000 in value” under 18 U.S.C. § 1030(c)(4)(A)(i)(I) because the unauthorized
28 access and collection of Private Information (i) caused a diminution in value of Plaintiffs’, Classes’,

1 and Subclasses' Private information, both of which occurred to millions of individuals, easily
2 aggregating at least \$5,000 in value.

3 326. For these reasons, and those discussed in this Complaint, Plaintiffs, Classes, and
4 Subclasses are entitled to "maintain a civil action against the violator to obtain compensatory
5 damages and injunctive relief or other equitable relief." 18 U.S.C. § 1030(g).

6 **COUNT THREE: VIOLATION OF THE CALIFORNIA INVASION OF PRIVACY ACT**
7 **("CIPA"), CAL. PENAL CODE § 631, et seq.**
8 **(on behalf of All Plaintiffs and the ChatGPT, ChatGPT API User, Microsoft User Classes**
9 **against Defendants)**

10 327. Plaintiffs hereby incorporate Paragraphs 1 through 266 as if fully stated herein.

11 328. The California Invasion of Privacy Act ("CIPA") is codified at Cal. Penal Code §§
12 630 to 638. The Act begins with its statement of purpose:

13 The Legislature hereby declares that advances in science and
14 technology have led to the development of new devices and techniques
15 for the purpose of eavesdropping upon private communications and that
16 the invasion of privacy resulting from the continual and increasing use
17 of such devices and techniques has created a serious threat to the free
18 exercise of personal liberties and cannot be tolerated in a free and
19 civilized society.

20 Cal. Penal Code § 630.

21 329. California Penal Code § 631(a) provides, in pertinent part:

22 Any person who, by means of any machine, instrument, or contrivance,
23 or in any other manner . . . willfully and without the consent of all parties
24 to the communication, or in any unauthorized manner, reads, or attempts
25 to read, or to learn the contents or meaning of any message, report, or
26 communication while the same is in transit or passing over any wire, line,
27 or cable, or is being sent from, or received at any place within this state;
28 or who uses, or attempts to use, in any manner, or for any purpose, or to
communicate in any way, any information so obtained, or who aids,
agrees with, employs, or conspires with any person or persons to lawfully
do, or permit, or cause to be done any of the acts or things mentioned
above in this section, is punishable by a fine not exceeding two thousand
five hundred dollars

329. California Penal Code § 632(a) provides, in pertinent part:

A person who, intentionally and without the consent of all parties to a
confidential communication, uses an electronic amplifying or recording
device to eavesdrop upon or record the confidential communication,

1 whether the communication is carried on among the parties in the presence
2 of one another or by means of a telegraph, telephone, or other device,
3 except a radio, shall be punished by a fine not exceeding two thousand five
4 hundred dollars

5 331. Under either section of the CIPA, a defendant must show it had the consent of all
6 parties to a communication.

7 332. OpenAI has its principal place of business in California; designed, contrived, and
8 effectuated its scheme to track users from California; and has adopted California substantive law to
9 govern its relationship with its users. Defendants conspired with OpenAI to effectuate these schemes
10 in and through California.

11 333. At all relevant times, Defendants' tracking and interceptions of the Plaintiffs' and
12 Class members' internet communications was without authorization and consent from the Plaintiffs,
13 Class members, and the websites they were browsing. The interception by Defendants was unlawful
14 and tortious.

15 334. Defendants' non-consensual tracking of the Plaintiffs' and Class members' internet
16 communications was designed to attempt to learn at least some meaning of the content in the URLs
17 and the communications that Plaintiffs and Class members were engaged in.

18 335. The following items constitute "machine[s], instrument[s], or contrivance[s]" under
19 the CIPA, and even if they do not, Google's deliberate and admittedly purposeful scheme that
20 facilitated its interceptions falls under the broad statutory catch-all category of "any other manner":

- 21 a. The computer codes and programs Defendants used to track the Plaintiffs' and
22 Class members' communications;
- 23 b. The Plaintiffs' and Class members' browsers and mobile applications;
- 24 c. The Plaintiffs' and Class members' computing and mobile devices;
- 25 d. Defendants' web and ad servers;
- 26 e. The web and ad-servers of websites from which Defendants tracked and
27 intercepted the Plaintiffs' and Class members' communications;
- 28 f. The computer codes and programs that Defendants used to effectuate tracking
and interception of the Plaintiffs' and Class members' communications; and

1 g. The plan Defendants carried out to effectuate the tracking and interception of
2 the Plaintiffs’ and Class members’ communications.

3 336. The data collected by Defendants constituted “confidential communications,” as that
4 term is used in Section 632, because Plaintiffs and Class members had objectively reasonable
5 expectations of privacy that the information would not be used for Defendants’ AI products.

6 337. Plaintiffs and Class members have suffered loss by reason of these violations,
7 including, but not limited to, violation of their rights to privacy and loss of value in their personally-
8 identifiable information.

9 338. Pursuant to California Penal Code § 637.2, Plaintiffs and Class members have been
10 injured by the violations of California Penal Code §§ 631 and 632, and each seek damages for the
11 greater of \$5,000 or three times the amount of actual damages, as well as injunctive relief

12
13 339. Plaintiffs bring this claim individually and on behalf of the members of the proposed
14 Class against Defendants.

15 340. CIPA § 631(a) imposes liability for “distinct and mutually independent patterns of
16 conduct.” *Tavernetti v. Superior Ct.*, 22 Cal. 3d 187, 192 (1978). Thus, to establish liability under
17 CIPA § 631(a), a plaintiff need only establish that the defendant, “by means of any machine,
18 instrument, contrivance, or in any other manner,” does any of the following:

19 Intentionally taps, or makes any unauthorized connection, whether
20 physically, electrically, acoustically, inductively or otherwise, with any
21 telegraph or telephone wire, line, cable, or instrument, including the wire,
line, cable, or instrument of any internal telephonic communication system,

22 **OR**

23 Willfully and without the consent of all parties to the communication, or in
24 any unauthorized manner, reads or attempts to read or learn the contents or
meaning of any message, report, or communication while the same is in
transit or passing over any wire, line or cable or is being sent from or
received at any place within this state,

25 **OR**

26 Uses, or attempts to use, in any manner, or for any purpose, or to
27 communicate in any way, any information so obtained,

28 **OR**

Aids, agrees with, employs, or conspires with any person or persons to unlawfully do,

1 or permit, or cause to be done any of the acts or things mentioned above in this section.
2 Cal. Penal Code § 631 (Deering 2023).

3 341. Section 631(a) is not limited to phone lines, but also applies to “new technologies”
4 such as computers, the Internet, and email. *See Matera v. Google Inc.*, No. 15-CV-04062-LHK,
5 2016 U.S. Dist. LEXIS 107918, at *61-*63 (N.D. Cal. Aug. 12, 2016) (CIPA applies to “new
6 technologies” and must be construed broadly to effectuate its remedial purpose of protecting
7 privacy); *Bradley v. Google, Inc.*, 2006 WL 3798134, at *5-6 (N.D. Cal. Dec. 22, 2006) (CIPA
8 governs “electronic communications”); *In re Facebook, Inc. Internet Tracking Litigation*, 956 F.3d
9 589, 598-99 (9th Cir. 2020) (reversing dismissal of CIPA and common law privacy claims based on
10 Facebook’s collection of consumers’ Internet browsing history).

11 342. Defendants’ ChatGPT platform is a “machine, instrument, contrivance, or ... other
12 manner” used to engage in the prohibited conduct at issue here.

13 **I. Defendants’ Interception of Communications of ChatGPT API Class Members**
14 **which occurred on Third-Party Websites, Platforms, Applications, Programs which**
15 **have integrated ChatGPT API. [Microsoft User Subclass is Excluded]**

16 343. The allegations for violation of CIPA § 631(a) arise out of Defendants’ interception
17 of Plaintiffs, ChatGPT API Class Members’ (collectively referred to as Chat-GPT API Class and
18 Subclass) communications which occurred on various applications, platforms, websites which
19 integrate ChatGPT technology (i.e., Stripe, Snapchat, etc.).

20 344. The transmissions of Plaintiffs’ and ChatGPT API Class Members’ communications
21 (including but not limited to chats, comments, replies, searches, keystrokes, mouse
22 clicks/movements, signals, browser activity, or other data, activity, or intelligence) on various
23 applications, programs, platforms, websites which integrate ChatGPT API (i.e., Stripe, Snapchat,
24 etc.) qualify as “electronic communications” under Cal. Penal Code §629.51(2).

25 345. By incorporating ChatGPT technology on third party platforms, Defendants are in the
26 unique position of having unrestricted, real-time access to the users’ every input, move, chat,
27 comment, reply, search, keystroke, or other browser activity/communication on the third-party
28 platform.

1 346. As Plaintiffs and ChatGPT API Class Members interact with the third-party platform,
2 Defendants intentionally tap, electrically or otherwise, the lines of internet communication between
3 Plaintiffs and ChatGPT API Class Members, and/or third-party entities.

4 347. In disregard for Plaintiffs’ and ChatGPT API Class Members’ privacy rights,
5 Defendants act as a third-party “eavesdropper”, redirecting Plaintiffs and Chat-GPT API Members’
6 electronic communications to Defendants’ own servers for appropriation, and training of their
7 Products.

8 348. Defendants’ interception of the contents of Plaintiffs’ and ChatGPT API Class
9 Members’ communications happens contemporaneously with their exchange of such
10 communications, whether such communications are directed to Plaintiffs’ and ChatGPT API Class
11 Members’ friends, colleagues, or third-party entities. As described above, the ChatGPT technology,
12 integrated on various platforms, is designed to simultaneously intercept and send a recording of
13 each keystroke, mouse click, movement, writing, or other data, activity, or intelligence to
14 Defendants sufficient to not only identify Plaintiffs and ChatGPT API Class Members’, but also to
15 be able to understand, collect, and use for training Plaintiffs’ and ChatGPT API Class Members’
16 communications.

17 349. Through this calculated scheme of using ChatGPT technology, integrated on various
18 non-ChatGPT platforms (such as Snapchat, Stripe etc.) to intercept, acquire, transmit, and record
19 Plaintiffs’ and ChatGPT API Class Members’ electronic communications, Defendants willfully and
20 without valid consent from all parties to the communication, take unauthorized measures to read
21 and understand the contents or meaning of the electronic communications of Plaintiffs and ChatGPT
22 API Class. The interception and recording of electronic communications occurs while the electronic
23 communications are in transit or passing over any wire, line, or cable, or are being sent from or
24 received at any place.

25 350. Plaintiffs and ChatGPT API Class Members did not authorize Defendants to acquire
26 the content of their communications for the purposes of training Defendants’ Products.

27 351. Plaintiffs, individually, on behalf of the GPT API Class, also seek all monetary and
28 non-monetary relief allowed by law, including actual damages, statutory damages in accordance

1 with § 637.2(a), punitive damages, preliminary and other equitable or declaratory relief, and
2 attorneys' fees and costs.

3 **II. Microsoft's Interception of ChatGPT User Class Members' Communications on**
4 **ChatGPT**

5 352. The allegations for violation of CIPA § 631(a) arise out of Defendant Microsoft's
6 interception of Plaintiffs' and ChatGPT User Class Members' communications which occurred on
7 ChatGPT platform.

8 353. The transmissions of Plaintiffs' and ChatGPT User Class Members' communications
9 (including but not limited to chats, comments, replies, searches, keystrokes, mouse
10 clicks/movements, signals, browser activity, or other data, activity, or intelligence) on ChatGPT
11 qualify as "electronic communications" under Cal. Penal Code §629.51(2).

12 354. By developing ChatGPT and controlling the extent of training/development of this
13 program, Defendants are in the unique position of having unrestricted, real-time access to the users'
14 every input, move, mouse click, chat, comment, reply, search, keystroke, browser activity, or other
15 data, activity, or intelligence on ChatGPT.

16 355. As Plaintiffs and ChatGPT User Class Members ask questions, or otherwise interact
17 with Defendant Open AI, Defendant Open AI intentionally aids and abets Defendant Microsoft to
18 intentionally tap and intercept, electrically or otherwise, the lines of internet communications of
19 Plaintiffs' and Chat-GPT User Class Members' searches and communications.

20 356. In disregard for Plaintiffs' and ChatGPT User Class Members' privacy rights,
21 Defendant Microsoft acts as a third-party "eavesdropper," redirecting Plaintiffs and Chat-GPT User
22 Class Members' electronic communications to Defendant Microsoft's own servers for
23 appropriation, and training of their Products.

24 357. Defendant Microsoft's interception of the contents of Plaintiffs' and ChatGPT User
25 Class Members' communications happens contemporaneously with their exchange of such
26 communications, whether such communications are directed to Defendant Open AI or third-party
27 entities (for instance, Expedia). As described above, the ChatGPT technology is designed to
28 simultaneously intercept and send a recording of each keystroke, mouse click, movement, writing,

1 or other data, activity, or intelligence to Defendant Microsoft sufficient to not only identify Plaintiffs
2 and Chat-GPT User Members, but also to be able to understand, collect, and use for training
3 Plaintiffs' and Chat-GPT User Class Members' communications.

4 358. Defendant Microsoft intercepted communications including all text entry input as a
5 search within ChatGPT as well as intercepted numerous other forms of a user's navigation and
6 interaction with ChatGPT.

7 359. Through this calculated scheme of using ChatGPT to intercept, acquire, transmit, and
8 record Plaintiffs' and ChatGPT User Class Members' electronic communications, Defendant
9 Microsoft willfully and without any iota of valid consent from all parties to the communication,
10 takes unauthorized measures to read and understand the contents or meaning of the electronic
11 communications of Plaintiffs and Chat-GPT User Class. The interception and recording of
12 electronic communications occur while the electronic communications are in transit or passing over
13 any wire, line, or cable, or are being sent from or received at any place.

14 360. In sending and in acquiring the content of Plaintiffs' and Class Members'
15 communications on ChatGPT, Defendants' purpose was tortious, and designed to violate federal
16 and state laws. By intentionally using, or endeavoring to use, the contents of the electronic
17 communications of Plaintiffs, ChatGPT User Class Members, while knowing or having reason to
18 know that the information was obtained through the interception of an electronic communication,
19 Defendant Microsoft violates CIPA § 631(a).

20 361. Additionally, under the fourth clause of §631(a), Defendant OpenAI aided, agreed
21 with, and conspired with Defendant Microsoft to accomplish the wrongful conduct at issue here.
22 *Graham v. Noom, Inc.*, 533 F. Supp. 3d 823, 831-32 (N.D. Cal. 2021) (while a party to a
23 communication may record the communication without triggering § 631(a) liability, it will be
24 subject to derivative liability where the third party is liable for recording the communications in
25 violation of the first, second or third clauses of § 631(a)); *Revitch v. New Moosejaw, LLC*, No. 18-
26 cv-06827-VC, 2019 WL 5485330, at *2 (N.D. Cal. 2019) (conversation participants may be liable
27 because § 631 "was designed to protect a person placing or receiving a call from a situation where
28 the person on the other end of the line permits an outsider to tap his telephone or listen in on the

1 call.”)

2 362. Plaintiffs, individually, on behalf of the GPT ChatGPT User Class Members, seek all
3 monetary and non-monetary relief allowed by law, including actual damages, statutory damages,
4 punitive damages, preliminary and other equitable or declaratory relief, and attorneys’ fees and
5 costs.

6 **III. Defendant Open AI’s Interception of Microsoft User Class Members which**
7 **occurred on Microsoft’s Websites, Platforms, Applications, Programs which have**
8 **integrated ChatGPT.**

9 363. The allegations for violation of CIPA § 631(a) arise out of Defendant Open AI’s
10 interception of Microsoft User Class Members’ (collectively “Microsoft Subclass”)
11 communications with their friends, family, colleagues, or other individuals or third-party entities,
12 which occurred on Microsoft platforms (Teams, Bing, Outlook etc.), which integrate ChatGPT API.

13 364. The transmissions of Plaintiffs’ and Microsoft Subclasses’ communications
14 (including but not limited to chats, comments, replies, searches, keystrokes, signals, mouse
15 clicks/movements, , browser activity, or other data, activity, or intelligence) on Microsoft’s various
16 applications, programs, platforms, websites which integrate ChatGPT API qualify as “electronic
17 communications” under Cal. Penal Code §629.51(2).

18 365. By integrating ChatGPT technology within the entire Microsoft suite, Defendant
19 OpenAI is in the unique position of having unrestricted, real-time access to the users’ every input,
20 move, mouse click, chat, comment, reply, search, keystroke, browser activity, or other data, activity,
21 or intelligence.

22 366. As Plaintiffs and Microsoft Subclasses interact with each other or the third-party
23 entities, Defendant OpenAI intentionally taps, electrically or otherwise intercept, the lines of
24 internet communications between Plaintiffs, Microsoft Subclasses, and/or third-party entities.

25 367. In disregard for Plaintiffs’ and Microsoft Subclasses Members’ privacy rights,
26 Defendant OpenAI acts as a third-party “eavesdropper,” redirecting Plaintiffs and Microsoft
27 Subclasses Members’ electronic communications to Defendants’ own servers for appropriation, and
28 training of their Products.

1 368. Defendant Open AI’s interception of the contents of Plaintiffs’ and Microsoft
2 Subclasses Members’ communications happens contemporaneously with their exchange of such
3 communications on Microsoft platforms, whether such communications are directed to Plaintiffs’
4 and Microsoft Subclasses Members’ friends, colleagues, or third-party entities. As described above,
5 the ChatGPT API is designed to simultaneously intercept and send a recording of each keystroke,
6 mouse click, signal, movement, writing, or other data, activity, or intelligence to Defendant Open
7 AI sufficient to not only identify Plaintiffs and Microsoft Subclasses Members, but also to be able
8 to understand, collect, and use for training Plaintiffs’ and Microsoft Subclasses Members’
9 communications.

10 369. Additionally, under the fourth clause of §631(a), Defendant Microsoft aided, agreed
11 with, and conspired with Defendant OpenAI to implement AI technology within its own platforms.
12 The incorporation of such technology shares users’ electronic communications with Microsoft
13 platforms with OpenAI in an effort to accomplish the wrongful conduct at issue here. *Graham v.*
14 *Noom, Inc.*, 533 F. Supp. 3d 823, 831-32 (N.D. Cal. 2021) (while a party to a communication may
15 record the communication without triggering § 631(a) liability, it will be subject to derivative
16 liability where the third party is liable for recording the communications in violation of the first,
17 second or third clauses of § 631(a)); *Revitch v. New Moosejaw, LLC*, No. 18-cv-06827-VC, 2019
18 WL 5485330, at *2 (N.D. Cal. 2019) (conversation participants may be liable because § 631 “was
19 designed to protect a person placing or receiving a call from a situation where the person on the
20 other end of the line permits an outsider to tap his telephone or listen in on the call.”)

21 370. Plaintiffs, individually, on behalf of the Microsoft Subclasses Members, seek all
22 monetary and non-monetary relief allowed by law, including actual damages, statutory damages,
23 punitive damages, preliminary and other equitable or declaratory relief, and attorneys’ fees and
24 costs.

25 371. Unless enjoined, Defendants will continue to commit the illegal acts alleged here.

26 372. Plaintiffs and Class Members seek all relief available under Cal. Penal Code § 637.2,
27 including injunctive relief and statutory damages of \$5,000 per violation.
28

1 **COUNT FOUR: VIOLATION OF CALIFORNIA UNFAIR COMPETITION LAW (Cal.**
2 **Bus. & Prof. Code §§ 17200, et seq.)**
3 **(on behalf of All Plaintiffs and the Classes against Defendants)**

4 373. Plaintiffs hereby incorporate Paragraphs 1 through 266 as if fully stated herein.

5 374. As discussed above, Plaintiffs believe that California law should apply to all
6 claimants, including out of state residents.

7 375. California Business & Professions Code, sections 17200, *et seq.* (the “UCL”) prohibits unfair competition and provides, in pertinent part, that “unfair competition shall mean and
8 include unlawful, unfair or fraudulent business practices and unfair, deceptive, untrue or misleading
9 advertising.”

10 **I. Unlawful**

11 376. Defendants engaged in and continue to engage in “unlawful” business acts and
12 practices under the Unfair Competition Law because Defendants took, accessed, intercepted,
13 tracked, collected, or used the Plaintiffs’ and Nationwide Classes’ Private Information, including
14 but not limited to their private conversations, personally identifiable information, financial and
15 medical data, keystrokes, searches, cookies, browser activity and other data, and shared this
16 information with each other, while also using this information to train Defendants’ AI Products.
17 Defendants’ unlawful conduct is as follows:

- 18 a) Web-Scraping and Interception of Communications, Private Information and Data:
19 Defendants scraped nearly the entire internet in order to train their AI Products, and
20 in this process, Defendants accessed, and stole private conversations, personal
21 information, and other private data from websites including Reddit, Twitter, TikTok,
22 Spotify, YouTube, and other websites, without consent of the individuals.
23 Defendants’ illegal web scraping violates privacy laws, and other laws outlined in this
24 complaint. Defendants failed to register as data brokers under California law as
25 required.
- 26 b) Defendants’ Intercepted Communications and Accessed, Collected, and Tracked
27 Private Information from Platforms Which Integrated ChatGPT: Defendants
28 intercepted, tracked, and recorded communications, messages, chats, web activity,

1 user activity, associated cookies, keystrokes and other Private Information through its
2 ChatGPT technology integrated within hundreds of applications (including but not
3 limited to Stripe, Snapchat, Expedia etc.) which were used to train Defendants’
4 Products. Defendants’ illegal tracking of such data, which is subsequently used to
5 train Defendants’ AI products violates privacy laws, California wiretapping law, and
6 other laws outlined in this complaint.

7 c) Open AI’s Interception of Communications and Accessed, Collected, and Tracked
8 Private Information on Microsoft Platforms: Defendant Microsoft aided Defendant
9 Open AI in intercepting, tracking, and recording communications, messages, chats,
10 web activity, user activity, associated cookies, and other Private Information through
11 its ChatGPT technology integrated within the entire Microsoft suite (Microsoft
12 Teams, Microsoft Outlook, Bing). Defendant’s Open AI illegal tracking of such data
13 and Defendant Microsoft’s aiding and abetting this conduct violates privacy laws,
14 California wiretapping law, and other laws outlined in this complaint.

15 d) Microsoft’s Interception of Communications and Accessed, Collected, and Tracked
16 Private Information on ChatGPT: Defendant OpenAI aided Defendant Microsoft in
17 intercepting, tracking, and recording communications, messages, chats, web activity,
18 user activity, associated cookies, and other Private Information by sharing access to
19 ChatGPT and sending all communications to Defendant Microsoft and its partners.

20 377. Defendants’ conduct as alleged herein was unfair within the meaning of the UCL. The
21 unfair prong of the UCL prohibits unfair business practices that either offend an established public
22 policy or that are immoral, unethical, oppressive, unscrupulous, or substantially injurious to
23 consumers.

24 378. Defendants’ conduct violates the EPCA, CFAA, CIPA, California Consumer Privacy
25 Act (“CCPA”), Cal. Civ. Code § 1798.100, *et seq.*, Section 5 of the Federal Trade Commission Act
26 (“FTCA”), Cal. Bus. & Prof. Code § 22575, *et seq.*, and other tort claims stated in this lawsuit. The
27 violations of EPCA, CFAA, CIPA, and other tort claims stated in this lawsuit, are incorporated
28 herein by reference.

1 379. Under the CCPA, a business that collects consumers’ personal information is
2 required, at or before the point of collection, to provide notice to consumers indicating: (1) “[t]he
3 categories of personal information to be collected and the purposes for which the categories of
4 personal information are collected or used and whether that information is sold or shared”; (2) “the
5 categories of sensitive personal information to be collected and the purposes for which the
6 categories of sensitive personal information are collected or used, and whether that information is
7 sold or shared.”; and (3) “[t]he length of time the business intends to retain each category of personal
8 information . . .” Cal. Civ. Code § 1798.100(a).

9 380. “Personal information” is defined by the CCPA as “information that identifies, relates
10 to, describes, is reasonably capable of being associated with, or could reasonably be linked, directly
11 or indirectly, with a particular consumer or household.” Cal. Civ. Code § 1798.140(v)(1).

12 381. As alleged, Defendant uses web scraping technology to collect information from
13 webpages across the internet and, in so doing, Defendant gathers and compiles personal information
14 about consumers that is reflected on those webpages.

15 382. Because Defendants conduct web scraping across millions of web pages, without
16 asking the affected consumers their permission to use their content for training, Defendants do not,
17 and cannot provide consumers with the notice required by Cal. Civ. Code § 1798.100(a) at or before
18 the point of collection. Similarly, when Defendants intercept and wiretap users’ communications on
19 various platforms which integrate ChatGPT, Microsoft platforms, and ChatGPT platforms, to use
20 these intercepted communications and gathered data to train their Products. Defendants never
21 notified Plaintiffs and affected Nationwide Classes Members of this extensive wiretapping, and
22 more importantly, that this information would be used for commercial purposes and development
23 of Defendants’ Products. Therefore, Defendants failed to provide notice to the affected consumers
24 as required by Cal. Civ. Code § 1798.100(a).

25 383. Defendant’s failure to provide notice to Plaintiffs and Nationwide Classes Members
26 whose personal information is collected through the process of web scraping and illegal wiretapping
27 is unlawful and violates Cal. Civ. Code § 1798.100(a).

28 384. The CCPA further grants consumers the right to “request that a business that collects

1 a consumer’s personal information disclose to that consumer the categories and specific pieces of
2 personal information the business has collected.” Cal. Civ. Code § 1798.100(b).

3 385. Upon receipt of a verifiable request for disclosure pursuant to Section 1798.110, a
4 business must “disclose any personal information it has collected about a consumer, directly or
5 indirectly, including through or by a service provider or contractor, to the consumer . . .” Cal. Civ.
6 Code § 1798.130 (3)(A).

7 386. Any disclosure must provide the requesting consumer with all of the following: (1)
8 “The categories of personal information it has collected about that consumer”; (2) “The categories
9 of sources from which the personal information is collected”; (3) “The business or commercial
10 purpose for collecting, selling, or sharing personal information” (4) “The categories of third parties
11 to whom the business discloses personal information”; and (5) “The specific pieces of personal
12 information it has collected about that consumer.” Cal. Civ. Code § 1798.110(a).

13 387. Consumers also “have the right to request that a business delete any personal
14 information about the consumer which the business has collected from the consumer.” Cal. Civ.
15 Code § 1798.105(a).

16 388. Pursuant to Cal. Civ. Code §§ 1798.100(b) and 1798.130(a), OpenAI’s privacy policy
17 provides a method by which California residents who have had their data collected may request
18 disclosure of the categories and specific pieces of personal information OpenAI has collected about
19 them.²³⁵ Open AI’s privacy policy specifically states that consumers “may have certain statutory
20 rights in relation to their Personal Information,” including the right to “Access your Personal
21 Information.”²³⁶

22 389. To exercise their right to access the Personal Information OpenAI has collected about
23 them, consumers are instructed to email their request for disclosure to dsar@openai.com.²³⁷

24 390. Under the heading “Additional U.S. State Disclosures,” the privacy policy states that
25 some users may have “[t]he right to know information about our processing of your Personal
26

27 ²³⁵ *Privacy Policy*, OPENAI, <https://openai.com/policies/privacy-policy> (last updated June 23,
2023).

28 ²³⁶ *Id.*

²³⁷ *Id.*

1 Information, including the specific pieces of Personal Information that we have collected from you
2 . . .”²³⁸ Users are instructed that, “to the extent applicable under local law, [they] can exercise privacy
3 rights. . . by submitting a request to dsar@openai.com.”²³⁹

4 391. Yet OpenAI fails to disclose that once its AI Products have been trained on an
5 individual’s information, that information has been included into the product and cannot reasonably
6 be extracted. Whether individuals’ information was collected through web scraping or obtained
7 through interception from ChatGPT, or other platforms incorporating ChatGPT, this information,
8 once used to train Products, cannot be extracted. Therefore, Defendants violated and continue to
9 violate CCPA.

10 392. Plaintiffs, individually and on behalf of the Nationwide Classes seek: (i) an injunction
11 requiring OpenAI to revise its privacy policy to fully disclose all information required under CCPA,
12 and to delete all information previously collected in violation of these laws; (ii) relief under Cal.
13 Bus. & Prof. Code § 17200, et seq., including, but not limited to, restitution to Plaintiffs and other
14 members of the Nationwide Classes of money or property Defendants acquired by means of their
15 unlawful business practices; and, as a result of bringing this action to vindicate and enforce an
16 important right affecting the public interest, (iii) reasonable attorney’s fees (pursuant to Cal. Code
17 of Civ. P. § 1021.5).

18 393. Defendants’ unlawful actions in violation of the UCL have caused and are likely to
19 cause substantial injury to consumers that consumers cannot reasonably avoid themselves and that
20 is not outweighed by countervailing benefits to consumers or competition.

21 394. As a direct and proximate result of Defendants’ misconduct, Plaintiffs and Nationwide
22 Classes Members had their private communications containing information related to their sensitive
23 and confidential Private Information intercepted, disclosed, and used by third parties, including but
24 not limited to each Defendant.

25 395. As a result of Defendants’ unlawful conduct, Plaintiffs and Nationwide Classes
26 Members suffered an injury, including violation to their rights of privacy, loss of value and privacy
27

28 ²³⁸ *Id.*

²³⁹ *Id.*

1 of their Private Information, loss of control over their sensitive personal information, and suffered
2 embarrassment and emotional distress as a result of this unauthorized scraping, interception,
3 sharing, and misuse of information.

4 **II. Unfair**

5 396. Defendants' conduct as alleged herein was unfair within the meaning of the UCL. The
6 unfair prong of the UCL prohibits unfair business practices that either offend an established public
7 policy or that are immoral, unethical, oppressive, unscrupulous or substantially injurious to
8 consumers.

9 397. Defendants also engaged in business acts or practices deemed "unfair" under the UCL
10 because, as alleged above, Defendants failed to disclose that they scraped information belonging to
11 millions of internet users without the users' consent. Defendants also failed to disclose that they
12 used the stolen information to train their Products, without consent of the internet users.
13 Furthermore, Defendants failed to disclose that they were intercepting, tracking Private Information
14 belonging to millions of ChatGPT users, and the users of other platforms which integrated ChatGPT.
15 Private Information obtained from individual uses of ChatGPT and other platforms which integrate
16 ChatGPT was and is continued to be used to train Defendants' Products, without consent of the
17 users.

18 398. Unfair acts under the UCL have been interpreted using three different tests: (1)
19 whether the public policy which is a predicate to a consumer unfair competition action under the
20 unfair prong of the UCL is tethered to specific constitutional, statutory, or regulatory provisions;
21 (2) whether the gravity of the harm to the consumer caused by the challenged business practice
22 outweighs the utility of the defendant's conduct; and (3) whether the consumer injury is substantial,
23 not outweighed by any countervailing benefits to consumers or competition, and is an injury that
24 consumers themselves could not reasonably have avoided.

25 399. Under the UCL, a business practice that is likely to deceive an ordinary consumer
26 constitutes a deceptive business practice. Defendants' conduct was deceptive in numerous respects.

27 400. Defendants' misrepresentations and omissions include both implicit and explicit
28 representations.

1 401. Defendant OpenAI represented, throughout the Class Period, that it would “respect
2 your privacy and [is] strongly committed to keeping secure any information we obtain from you or
3 about you.”

4 402. Defendants’ conduct, as alleged herein, was fraudulent within the meaning of the
5 UCL. Defendants made deceptive misrepresentations and omitted known material facts in
6 connection with the solicitation, interception, disclosure, and use of Plaintiffs’ and Class Members’
7 User Data. Defendants actively concealed and continued to assert misleading statements regarding
8 their protection and limitation on the use of the User Data. Meanwhile, Defendants were collecting
9 and sharing Plaintiffs’ and Class Members’ User Data without their authorization or knowledge in
10 order to profit off of the information, and to deliver advertisements to Plaintiffs and Class Members,
11 among other unlawful purposes.

12 403. Defendants’ conduct, as alleged herein, was unlawful within the meaning of the UCL
13 because Defendants violated regulations and laws as discussed herein, including but not limited to
14 HIPAA, Section 5 of the Federal Trade Commission Act (“FTCA”), 15 U.S.C. § 45 and the CIPA.

15 404. Defendants reaped profits from these actions in the form of increased company
16 valuation, investments, improved language model performance, and dominance in the AI field.

17 405. Defendants’ unlawful actions in violation of the UCL have caused and are likely to
18 cause substantial injury to consumers that consumers cannot reasonably avoid themselves and that
19 is not outweighed by countervailing benefits to consumers or competition.

20 406. As a direct and proximate result of Defendants’ misconduct, Plaintiffs and Nationwide
21 Classes Members had their private communications containing information related to their sensitive
22 and confidential User Data intercepted, disclosed, and used by third parties, including but not limited
23 to each Defendant.

24 407. As a result of Defendants’ unlawful conduct, Plaintiffs and Nationwide Classes
25 Members suffered an injury, including violation to their rights of privacy, loss of the privacy of their
26 PHI/PII, loss of control over their sensitive personal information, and suffered aggravation,
27 inconvenience, and emotional distress.

28 408. Further, Defendants’ conduct is immoral, unethical, oppressive, unscrupulous and

1 substantially injurious to Plaintiffs and Nationwide Classes Members, and there are no greater
2 countervailing benefits to consumers or competition.

3 409. Plaintiffs, as well as the Nationwide Classes Members, were harmed by Defendants'
4 violations of Cal. Bus. & Prof. Code §17200. Defendants' practices were a substantial factor and
5 caused injury in fact and actual damages to Plaintiffs and Nationwide Classes Members.

6 410. As a direct and proximate result of Defendants' deceptive acts and practices, Plaintiffs
7 and Nationwide Classes Members have suffered and will continue to suffer an ascertainable loss of
8 money or property, real or personal, and monetary and non-monetary damages, as described above,
9 including the loss or diminishment in value of their Private Information and the loss of the ability
10 to control the use of their Private Information, which allowed Defendants to profit at the expense of
11 Plaintiffs and Nationwide Classes Members.

12 411. Plaintiffs' and Nationwide Classes Members' Personal Information has tangible
13 value; it is now in the possession of Defendants, who has used and will continue to use it for financial
14 gain.

15 412. Plaintiffs' and Nationwide Classes Members injury was the direct and proximate
16 result of Defendants' conduct described herein.

17 413. Defendants' retention of Plaintiffs' and Nationwide Classes Members' Personal
18 Information presents a continuing risk to them as well as the general public.

19 414. Plaintiffs, individually and on behalf of the Nationwide Classes Members, seek: (1)
20 an injunction requiring Defendants to permanently delete, destroy or otherwise sequester the Private
21 Information collected without consent; (2) compensatory restitution of Plaintiffs' and Nationwide
22 Classes Members money and property lost as a result of Defendants' acts of unfair competition; (3)
23 disgorgement of Defendants' unjust gains; and (4) reasonable attorney's fees (pursuant to Cal. Code
24 of Civ. Proc. § 1021.5).

25 415. Had Plaintiffs and Nationwide Classes Members known Defendants would disclose
26 and misuse their User Data in contravention of Defendants' representations, they would not have
27 used Defendants' Products.

28 416. Defendants' unlawful actions in violation of the UCL have caused and are likely to

1 cause substantial injury to consumers that consumers cannot reasonably avoid themselves and that
2 is not outweighed by countervailing benefits to consumers or competition.

3 417. As a direct and proximate result of Defendants’ misconduct, Plaintiffs and Nationwide
4 Classes Members had their private communications containing information related to their sensitive
5 and confidential Private Information intercepted, disclosed, and used by Defendants, to train their
6 Products.

7 418. As a result of Defendants’ unlawful conduct, Plaintiffs and Nationwide Classes
8 Members suffered an injury, including violation to their rights of privacy, loss of the privacy of their
9 Private Information loss of control over their sensitive personal information, and suffered
10 aggravation, inconvenience, and emotional distress.

11 **COUNT FIVE: NEGLIGENCE**

12 **(on behalf of All Plaintiffs and the Classes against Defendants)**

13 419. Plaintiffs hereby incorporate Paragraphs 1 through 266 as if fully stated herein.

14 420. Defendants owed a duty to Plaintiffs and the Classes to exercise due care in: (a)
15 obtaining data to train their Products; (b) not using individual’s private information to train
16 Defendants’ AI; (c) ensuring that individuals’ private data is not shared with or disclosed to
17 unauthorized parties (including Defendant Microsoft); (d) destroying personal information to which
18 Defendants had no legal right to possess.

19 421. Defendants’ duties to use reasonable care arose from several sources, including those
20 described below. Defendants had a common law duty to prevent foreseeable harm to others,
21 including Plaintiffs and members of the Classes, who were the foreseeable and probable victims of
22 Defendants’ unlawful practices. Defendants acknowledge the Products are inherently unpredictable
23 and may even evolve to act against human interests. Nevertheless, Defendants collected and
24 continue to collect Private Information of millions of individuals and permanently feed the data to
25 the Products, to train the Products for Defendants’ commercial benefit. Defendants knowingly put
26 Plaintiffs and the Classes in a zone of risk that is incalculable – but unacceptable by any measure of
27 responsible data protection and use.

28 422. Defendants’ conduct as described above constituted an unlawful breach of their duty

1 to exercise due care in collecting, storing, and safeguarding Plaintiffs’ and the Classes Members’
2 Private Information by failing to protect this information.

3 423. Plaintiffs and Classes Members trusted Defendants to act reasonably, as a reasonably
4 prudent manufacturer of AI products, and also trusted Defendants not to use individuals’ Private
5 Information to train their AI products. Defendants failed to do so, and breached their duty.

6 424. Defendants’ negligence was, at least, a substantial factor in causing the Plaintiffs and
7 the Classes’ Private Information to be improperly accessed, disclosed, used for development and
8 training of a dangerous product, and in causing the Class members’ injuries.

9 425. The damages suffered by Plaintiffs and the Classes’ members was the direct and
10 reasonably foreseeable result of Defendants’ negligent breach of their duties to adequately design,
11 implement, and maintain reasonable practices to (a) avoid web scraping without consent of the
12 users; (b) avoid using Personal Information to train their AI products; and (c) avoid collecting and
13 sharing Users’ data with each other.

14 450. Defendants’ negligence directly caused significant harm to Plaintiffs and the Classes.

15 **COUNT SIX: INVASION OF PRIVACY**
16 **(on behalf of All Plaintiffs and the Classes against Defendants)**

17 426. Plaintiffs hereby incorporate Paragraphs 1 through 266 as if fully stated herein.

18 427. The right to privacy in California’s Constitution creates a right of action against
19 private entities such as Defendants.

20 428. Plaintiffs’ and Class members’ expectation of privacy is deeply enshrined in
21 California’s Constitution. Article I, section 1 of the California Constitution provides: “All people
22 are by nature free and independent and have inalienable rights. Among these are enjoying and
23 defending life and liberty, acquiring, possessing, and protecting property and pursuing and obtaining
24 safety, happiness, *and privacy*.” (Emphasis added).

25 429. The phrase “and privacy” was added in 1972 after voters approved a proposed
26 legislative constitutional amendment designated as Proposition 11. Critically, the argument in favor
27 of Proposition 11 reveals that the legislative intent was to curb businesses’ control over the
28 unauthorized collections and use of consumers’ personal information, stating:

The right of privacy is the right to be left alone...It prevents government

1 and business interests from collecting and stockpiling unnecessary
2 information about us and from misusing information gathered for one
3 purpose in order to serve other purposes or to embarrass us.
4 Fundamental to our privacy is the ability to control circulation of
5 personal information. This is essential to social relationships and
6 personal freedom.

7 430. The principal purpose of this constitutional right was to protect against unnecessary
8 information gathering, use, and dissemination by public and private entities, including Defendants.

9 431. To plead a California constitutional privacy claim, a plaintiff must show an invasion
10 of: 1) a legally protected privacy interest; 2) where the plaintiff had a reasonable expectation of
11 privacy in the circumstances; and 3) conduct by the defendant constituting a serious invasion of
12 privacy.

13 432. As described herein, Defendants have intruded upon the following legally protected
14 privacy interests:

- 15 a. The Federal Wiretap Act as alleged herein;
- 16 b. The California Wiretap Act as alleged herein;
- 17 c. A Fourth Amendment right to privacy contained on personal computing
18 devices, including web-browsing activity, as explained by the United States
19 Supreme Court in the unanimous decision of *Riley v. California*;
- 20 d. The California Constitution, which guarantees Californians the right to
21 privacy; and
- 22 e. Defendant's Privacy Policies and policies referenced therein.

23 433. Plaintiffs and Class members had a reasonable expectation of privacy under the
24 circumstances in that Plaintiffs and Class members could not reasonably expect Defendants would
25 commit acts in violation of federal and state civil and criminal laws.

26 434. Defendant's actions constituted a serious invasion of privacy in that it:

- 27 a. Invaded a zone of privacy protected by the Fourth Amendment, namely the
28 right to privacy in data contained on personal computing devices, including
web search and browsing histories;
- b. Violated several federal criminal laws, including the Wiretap Act;
- c. Violated dozens of state criminal laws on wiretapping and invasion of privacy,

- 1 including the California Invasion of Privacy Act;
- 2 d. Invaded the privacy rights of hundreds of millions of Americans (including
- 3 Plaintiffs and class members) without their consent;
- 4 e. Constituted the unauthorized taking of valuable information from hundreds of
- 5 millions of Americans through deceit; and
- 6 f. Further violated Plaintiffs' and Class members' reasonable expectation of
- 7 privacy via Defendants' review, analysis, and subsequent uses of Plaintiffs'
- 8 and Class members' browsing activity that Plaintiffs and Class members
- 9 considered sensitive and confidential, and did not intend to be used in
- 10 Defendants' AI products.

11 435. Committing criminal acts against hundreds of millions of Americans constitutes an
12 egregious breach of social norms that is highly offensive.

13 436. The surreptitious and unauthorized tracking of the internet communications of
14 millions of Americans constitutes an egregious breach of social norms that is highly offensive.

15 437. Defendants' intentional intrusion into Plaintiffs' and Class members' internet
16 communications and their computing devices and web-browsers was highly offensive to a
17 reasonable person in that Google violated federal and state criminal and civil laws designed to
18 protect individual privacy and against theft.

19 438. The taking of personally-identifiable information from hundreds of millions of
20 Americans through deceit is highly offensive behavior.

21 439. Secret monitoring of web browsing is highly offensive behavior.

22 440. Following Defendants' unauthorized interception of the sensitive and valuable
23 personal information, the subsequent analysis and use of that activity to develop and refine
24 Defendants' AI products violated Plaintiffs' and Class Members' reasonable expectations of
25 privacy.

26 441. Wiretapping and surreptitious recording of communications is highly offensive
27 behavior.

28

1 442. Defendants’ lacked any legitimate business interest in tracking users then using that
2 information in AI products without their consent.

3 443. Plaintiffs and Class members have been damaged by Defendants’ invasion of
4 their privacy and are entitled to just compensation and injunctive relief.

5 **COUNT SEVEN: INTRUSION UPON SECLUSION**
6 **(on behalf of All Plaintiffs and the Classes against Defendants)**

7 444. Plaintiffs hereby incorporate Paragraphs 1 through 266 as if fully stated herein.

8 445. Plaintiffs asserting a claim for intrusion upon seclusion must plead: 1) intrusion into
9 a private place, conversation, or matter; and 2) in a manner highly offensive to a reasonable person.

10 446. In carrying out their scheme to track and intercept Plaintiffs’ and Class members’
11 communications and other private data, Defendants intentionally intruded upon the Plaintiffs’ and
12 Class members’ solitude or seclusion in that Defendants effectively placed themselves in the middle
13 of conversations to which they were not authorized parties.

14 447. Defendants’ actions were not authorized by Plaintiffs and Class members, the
15 Websites with which they were communicating, or the devices that Plaintiffs and Class members
16 were using to facilitate those communications.

17 448. Defendants’ intentional intrusion into those communications and Plaintiffs’ and Class
18 members’ devices was highly offensive to a reasonable person in that they violated federal and state
19 criminal and civil laws designed to protect individual privacy and against theft.

20 449. The taking of personally identifiable information from the hundreds of millions of
21 Americans through deceit is highly offensive behavior.

22 450. Defendants’ secret monitoring of web browsing is also highly offensive behavior.

23 451. Wiretapping and surreptitious recording of communications is highly offensive
24 behavior.

25 452. Public polling on internet tracking has consistently revealed that the overwhelming
26 majority of Americans believe it is important—or very important—to be “in control of who can get
27 information” about them; to not be tracked without their consent; and to be in “control[] of what
28 information is collected about [them].” This desire to control one’s information is especially
heightened in today’s electronic age and with the proliferation of AI products.

1 453. Plaintiffs and Class members have been damaged by Google’s invasion of their
 2 privacy and are entitled to reasonable compensation including but not limited to disgorgement of
 3 profits related to the unlawful internet tracking, collection, and use of their data in AI products.

4 **COUNT EIGHT: LARCENY/RECEIPT OF STOLEN PROPERTY CAL. PENAL CODE**

5 **§§ 496(a) and 496(c)**

6 **(on behalf of All Plaintiffs and the Classes against Defendants)**

7 454. Plaintiffs hereby incorporate Paragraphs 1 through 266 as if fully stated herein.

8 455. Courts recognize that internet users have a property interest in their personal
 9 information and data. *See Calhoun v. Google, LLC*, 526 F. Supp. 3d 605, 635 (N.D. Cal. 2021)
 10 (recognizing property interest in personal information and rejecting Google’s argument that “the
 11 personal information that Google allegedly stole is not property”); *In re Experian Data Breach*
 12 *Litigation*, SACV 15-1592 AG (DFMx), 2016 U.S. Dist. LEXIS 184500, at *14 (C.D. Cal. Dec. 29,
 13 2016) (loss of value of PII is a viable damages theory); *In re Marriott Int’l Inc. Customer Data Sec.*
 14 *Breach Litig.*, 440 F. Supp. 3d 447, 460-61 (D. Md. 2020) (“The growing trend across courts that
 15 have considered this issue is to recognize the lost property value of this [personal] information.”);
 16 *Simona Opris v. Sincera*, No. 21-3072, 2022 U.S. Dist. LEXIS 94192, at *20 (E.D. Pa. May 23,
 2022) (collecting cases).

17 456. Defendants owned and operated their AI Products and GPT Platforms (ChatGPT,
 18 ChatGPT Plug-Ins, ChatGPT API). Defendants illegally obtained vast amounts of private
 19 information to train their AI Products.

20 **A. Defendants’ Taking of Individual’s Private Information to Train Their AI**

21 **Violated Plaintiffs’ Property Interests**

22 457. Penal Code § 496(a) creates an action against “any” person who (1) receives “any”
 23 property that has been stolen or obtained in any manner constituting theft, knowing the property to
 24 be stolen or obtained, or (2) conceals, sells, withholds, or aids in concealing or withholding “any”
 25 property from the owner, knowing the property to be so stolen or illegally obtained.

26 458. Under Penal Code § 1.07(a)(38), “person” means “an individual, corporation, or
 27 association.” Thus, Defendants are persons under section 496(a).

28 459. As discussed above, Defendants stole the contents of the internet – everything

1 individuals posted, information about the individuals, personal data, medical information, and other
2 information – all used to create their Products to generate massive profits. At no point did
3 Defendants have individuals consent to take/scrape this information in order to train their AI
4 Products. Defendants meet the grounds for liability under Cal. Penal Code 496(a) because each of
5 them:

- 6 a. Knew that the taken information was stolen or obtained by theft, and with such
7 knowledge;
- 8 b. Concealed, withheld, or aided in concealing or withholding said data from their
9 rightful owners by unlawfully using the data to train their Products;
- 10 c. Defendants moved the data from the internet in order to feed it into their Products for
11 training.

12 460. Pursuant to California Penal Code 496(c), Plaintiffs, on behalf of themselves and the
13 Nationwide Class, seek actual damages, treble damages, costs of suit, and reasonable attorneys’
14 fees.

15 **B. Tracking, Collecting, and Sharing Private Information Without Consent**

16 461. As described above, in violation of Cal. Penal Code § 496(a) and (c), Defendants
17 unlawfully collected, used, and exercised dominion and control of Private Information belonging to
18 Plaintiffs and Classes Members.

19 462. Defendants wrongfully took Plaintiffs’, ChatGpt User Class’, ChatGPT API User
20 Class’, and Microsoft User Class’ (collectively “User Classes”) Private Information to be used to
21 feed into Defendants’ AI Products, to train and develop a dangerous technology.

22 463. Plaintiffs and the User Classes Members did not consent to such taking and misuse of
23 their personal data, and Private Information.

24 464. Defendants did not have consent from any state or local government agency allowing
25 them to engage in such taking and misuse of Private Information.

26 465. Defendants’ taking of Private Information was intended to deprive the owners of such
27 information from ability to use their Private Information in the way they chose.

28 466. Defendants did so to maximize their profits and become rich at the expense of

1 Plaintiffs and the Classes.

2 467. Defendants collected data allows Defendants and their AI to learn the unique patterns
3 of each individuals, their online activities, habits, and speech/writing patterns.

4 468. Defendants moved Private Information to store and collect it on Defendant
5 Microsoft's servers, and thereafter, feed it to their AI products.

6 469. As a result of Defendants' actions, Plaintiffs and User Classes Members seek
7 injunctive relief, in the form of Defendants' cessation of tracking practices in violation of state law,
8 and destruction of all personal data obtained in violation of state law.

9 470. As a result of Defendants' actions, Plaintiffs, Nationwide Classes, and User Classes
10 seek nominal, actual, treble, and punitive damages in an amount to be determined at trial. Plaintiffs,
11 Nationwide Classes, and User Classes seek treble and punitive damages because Defendants'
12 actions—which were malicious, oppressive, willful—were calculated to injure Plaintiffs and made
13 in conscious disregard of Plaintiffs' rights. Punitive damages are warranted to deter Defendants
14 from engaging in future misconduct.

15 471. Plaintiffs seek restitution for the unjust enrichment obtained by Defendants as a result
16 of the commercialization of Plaintiffs', Nationwide Classes', and User Classes' sensitive data.

17 **COUNT NINE: CONVERSION**

18 **(on behalf of All Plaintiffs and the Classes against Defendants)**

19 451. Plaintiffs hereby incorporate Paragraphs 1 through 266 as if fully stated herein.

20 472. The Nationwide Classes repeat and incorporate by reference all preceding paragraphs
21 as if fully set forth herein.

22 473. Property is the right of any person to possess, use, enjoy, or dispose of a thing,
23 including intangible things such as data or communications. Plaintiffs' and Nationwide Classes
24 Members' personal information is their property. *Calhoun v. Google LLC*, 526 F. Supp. 3d 605,
25 636 (N.D. Cal. 2021).

26 474. As described in the cause of action for Larceny / Receipt of Stolen Property, Cal.
27 Penal Code § 496(a) and (c), Defendants unlawfully collected, used, and exercised dominion and
28 control over the Nationwide Classes Members' personal and private information without

1 authorization.

2 475. Defendants wrongfully exercised control over Plaintiffs' and Nationwide Classes'
3 information and have not returned it.

4 476. Plaintiffs and Nationwide Classes Members have been damaged as a result of
5 Defendants' unlawful conversion of their property.

6 **COUNT TEN: UNJUST ENRICHMENT**
7 **(on behalf of All Plaintiffs and the Classes against Defendants)**

8 477. Plaintiffs hereby incorporate Paragraphs 1 through 266 as if fully stated herein.

9 478. By virtue of the unlawful, unfair and deceptive conduct alleged herein, Defendants
10 knowingly realized hundreds of millions of dollars in revenue from the use of the Personal
11 Information of Plaintiffs and Nationwide Classes Members for the commercial training of its
12 ChatGPT and other AI language models.

13 479. This Private and Personal Information, the value of the Private and Personal
14 Information, and/or the attendant revenue, were monetary benefits conferred upon Defendants by
15 Plaintiffs and the members of the Nationwide Classes.

16 480. As a result of Defendants' conduct, Plaintiffs and Nationwide Classes Members
17 suffered actual damages in the loss of value of their Private Information and the lost profits from
18 the use of their Private Information.

19 481. It would be inequitable and unjust to permit Defendants to retain the enormous
20 economic benefits (financial and otherwise) it has obtained from and/or at the expense of Plaintiffs
21 and Classes Members.

22 482. Defendants will be unjustly enriched if they are permitted to retain the economic
23 benefits conferred upon them by Plaintiffs and Nationwide Classes Members through Defendants'
24 obtaining the Private Information and the value thereof, and profiting from the unlawful,
25 unauthorized, and impermissible use of the Private Information of Plaintiffs and Nationwide Classes
26 members.

27 483. Plaintiffs and Nationwide Classes members are therefore entitled to recover the
28 amounts realized by Defendants at the expense of Plaintiffs and Nationwide Classes Members.

1 484. Plaintiffs and the Nationwide Classes have no adequate remedy at law.

2 485. Plaintiffs and the members of the Nationwide Classes are entitled to restitution,
3 disgorgement, and/or the imposition of a constructive trust to recover the amount of Defendants'
4 ill-gotten gains, and/or other sums as may be just and equitable.

5 **COUNT ELEVEN: NEW YORK GENERAL BUSINESS LAW §§ 349, et seq.**
6 **(on behalf of Plaintiff J.H. and the New York Subclasses against Defendants)**

7 486. Plaintiff J.H., individually and on behalf of the New York Subclasses, hereby
8 incorporates Paragraphs 1 through 266 as if fully stated herein.

9 487. Defendants engaged in deceptive acts or practices in the conduct of its business, trade,
10 and commerce or furnishing of services, in violation of N.Y. Gen. Bus. Law § 349, including:

- 11 a) Defendants have exploited Non-Users and Users of their Products, by stealing
12 such individuals' data at scale from web crawler caches without permission from
13 the data owners and without any way of segregating out any given Non-Users'
14 or User data from the datasets used to train OpenAI's LLMs upon request of
15 such individuals.
- 16 b) Defendants knew that they were collecting and/or profiting from individuals'
17 Personal Information and that the risk of collecting of such Personal
18 Information was highly likely. Defendants' actions in engaging in the above-
19 named deceptive acts and practices were negligent, knowing and willful,
20 and/or wanton and reckless with respect to the rights of Plaintiff J.H. and
21 members of the New York Subclasses;
- 22 c) As described herein, Defendants are misrepresenting that they have and are
23 complying with common law and statutory duties pertaining to the security and
24 privacy of Plaintiff J.H.'s and Subclass Members' data, including but not limited
25 to duties imposed by the FTC Act, 15 U.S.C. § 45 and N.Y. Gen. Bus. Law §§
26 349, et seq.
- 27 d) As described herein, Defendants have and are omitting, suppressing, and
28 concealing the material fact that they are stealing and profiting from the mass

1 collection and analysis of Plaintiff J.H.’s and New York Subclasses Members’
2 data at scale and without adequate or effective consent; and
3 e) Omitting, suppressing, and concealing the material fact that they did not comply
4 with common law and statutory duties pertaining to the security and privacy of
5 Plaintiff J.H.’s and Subclasses Members’ data, including but not limited to the
6 fact that they are functionally unable to delete such data once it has been
7 incorporated into their LLMs as training data.

8 488. Defendants’ representations and omissions were material because they were likely to
9 deceive reasonable consumers about the terms of use of their products, as well as the available
10 mechanisms for seeking to exert control over Plaintiff J.H.’s and New York Subclasses Members’
11 data.

12 489. Defendants acted intentionally, knowingly, and maliciously to violate New York’s
13 General Business Law, and recklessly disregarded the Plaintiff J.H.’s and New York Subclasses
14 Members’ rights.

15 490. As a direct and proximate result of Defendants’ deceptive and unlawful acts and
16 practices, Plaintiff J.H. and New York Subclasses Members have suffered and will continue to suffer
17 injury, ascertainable losses of money or property, and monetary and non-monetary damages, as
18 described herein.

19 491. Defendants’ deceptive and unlawful acts and practices complained of herein affected
20 the public interest and consumers at large, including millions of New Yorker User Class Members
21 and Non-User Subclass Members.

22 492. The above deceptive and unlawful practices and acts by Defendants caused substantial
23 injury to Plaintiff J.H. and New York Subclasses Members that they could not reasonably avoid.

24 493. Plaintiff J.H. and New York Subclasses Members seek all monetary and non-
25 monetary relief allowed by law, including actual damages or statutory damages of \$50 (whichever
26 is greater), treble damages, injunctive relief, and attorney’s fees and costs.

27 **PRAYER FOR RELIEF**

28 WHEREFORE, Plaintiffs on behalf of themselves and the Proposed Classes that they seek

1 to represent, respectfully requests the following relief:

- 2 A. Certify this action as a class action pursuant to Rule 23 of the Federal Rules of Civil
3 Procedure
- 4 B. Appoints Plaintiffs to represent the Classes;
- 5 C. Appoint undersigned counsel to represent the Classes;
- 6 D. Award compensatory damages (including treble damages, where appropriate) to
7 Plaintiffs and the Class members against Defendants for all damages sustained as a
8 result of Defendants' wrongdoing, in an amount to be proven at trial, including interest
9 thereon;
- 10 E. Award statutory (including treble damages, where appropriate) damages to Plaintiffs
11 and the Class members against Defendants;
- 12 F. Award nominal damages to Plaintiffs and the Class members against Defendants;
- 13 G. Non-restitutionary disgorgement of all profits that were derived, in whole or in part,
14 from Defendants' conduct;
- 15 H. Award punitive damages to Plaintiffs and the Class members against Defendants;
- 16 I. For all Counts, permanently restrain Defendants, and its officers, agents, servants,
17 employees, and attorneys, from the conduct at issue in this Action and otherwise
18 violating its policies with consumers, and award all other appropriate injunctive and
19 equitable relief deemed just and proper, including:
 - 20 1. Establishment of an independent body of thought leaders (the "AI
21 Council") who shall be responsible for approving uses of the Products
22 before, not after, the Products are deployed for said uses;
 - 23 2. Implementation of Accountability Protocols that hold Defendants
24 responsible for Product actions and outputs and barred from further
25 commercial deployment absent the Products' ability to follow a code of
26 human-like ethical principles and guidelines and respect for human
27 values and rights, and until Plaintiffs and Class Members are fairly
28 compensated for the stolen data on which the Products depend;

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28

3. Implementation of effective cybersecurity safeguards of the Products as determined by the AI Council, including adequate protocols and practices to protect Users’ PHI/PII collected through Users’ inputting such information within the Products as well as through Defendants’ massive web scraping, consistent with the industry standards, applicable regulations, and federal, state, and/or local laws;
4. Implementation of Appropriate Transparency Protocols requiring Defendants to clearly and precisely disclose the data they are collecting, including where and from whom, in clear and conspicuous policy documents that are explicit about how this information is to be stored, handled, protected, and used;
5. Requiring Defendants to allow Product users and everyday internet users to opt out of all data collection and stop the illegal taking of internet data, delete (or compensate for) any ill-gotten data, or the algorithms which were built on the stolen data;
6. Requiring Defendants to add technological safety measures to the Products that will prevent the technology from surpassing human intelligence and harming others;
7. Requiring Defendants to implement, maintain, regularly review and revise as necessary, a threat management program designed to appropriately monitor Defendants’ information networks for threats, both internal and external, and assess whether monitoring tools are appropriately configured, tested, and updated;
8. Establishment of a monetary fund (the “AI Monetary Fund” or “AIMF”) to compensate class members for Defendants’ past and ongoing misconduct to be funded by a percentage of gross revenues from the Products;
9. Appointment of a third-party administrator (the “AIMF Administrator”)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28

to administer the AIMF to members of the class as “data dividends” as fair and just compensation for the stolen data on which the Products depend;

10. Confirmation that Defendants have deleted, destroyed, and purged the PII/PHI of all relevant class members unless Defendants can provide reasonable justification for the retention and use of such information when weighed against the privacy interests of class members; and

11. Requiring all further and just corrective action, consistent with permissible law and pursuant to only those causes of action so permitted;

J. Award Plaintiffs and the Class members their reasonable costs and expenses incurred in this Action, including attorneys’ fees, costs, and expenses; and

K. Grant Plaintiffs and the Class Members such further relief as the Court deems appropriate.

JURY TRIAL DEMANDED

Plaintiffs demand a jury trial on all triable issues.

DATED: September 5, 2023

/s/ Michael F. Ram

MORGAN & MORGAN
COMPLEX LITIGATION GROUP
Michael F. Ram
John A. Yanchunis
Ryan J. McGee